Aalto University
School of Science
Master's Programme in Computer, Communication and Information Sciences

Nagadivya Balasubramaniam

Using Ethical Guidelines for Defining Critical Quality Requirements of AI Solutions

Master's Thesis
Espoo, September 30, 2019

Supervisor: Professor Marjo Kauppinen, Aalto University
Advisor: Dr. Kari Hiekkanen, Aalto University

Aalto University
School of Science
Master's Programme in
Computer, Communication and Information Sciences

ABSTRACT OF
MASTER'S THESIS

Author: Nagadivya Balasubramaniam
Title: Using Ethical Guidelines for Defining Critical Quality Requirements of AI Solutions

Date: September 30, 2019 **Pages:** 70

Professorship: Software and Service Engineering **Code:** SCI3043

Supervisor: Professor Marjo Kauppinen
Instructor: Dr. Kari Hiekkanen

Artificial Intelligence (AI) has become a fast-growing trend and an increasing number of companies are interested in developing AI solutions. Many companies have realised that black box nature of AI technologies can raise ethical questions. And they are defining ethical guidelines which can be utilised when developing AI solutions. The goal of this thesis was to explore how companies can solve potential ethical issues of AI solutions during requirements engineering activities. The main research method used in this thesis was a multiple-case study approach. Semi-structured interviews and document analysis were carried out to discover the AI ethical guidelines in three Finnish companies. The results are based on the open coding of the interview data and the ethical guidelines of the case companies.

The ethical guidelines developed by the case companies focus on solving potential ethical issues like accountability, fairness, privacy, safety, security, transparency and trust. All three case companies devised ethical guideline relating to fairness, transparency and trust. To capture different views on critical ethical issues, two of the case companies recommend using multi-disciplinary development teams consisting of experts like data scientists, designers, and sociologists. The results of this thesis suggest that companies can utilise the ethical guidelines of AI solutions to identify and prioritise the essential quality requirements of AI solutions.

Keywords:
AI ethical issues, Quality requirements, Ethical guidelines of AI, Requirements Engineering (RE)

Language: English

Acknowledgements

Writing this thesis has been fascinating and extremely rewarding. It is now time to thank everyone who has supported me in this Master's degree journey.

Foremost, I express my sincere and deepest gratitude to my supervisor Prof. Marjo Kauppinen for always supporting and motivating me to give my best. Thanks for her invaluable guidance, affectionate attitude and healthy criticism. Without her continuous encouragement it would have been not possible to complete this thesis.

I owe special thanks to my thesis advisor Dr. Kari Hiekkanen who was always ready to support and encourage on my research. I am obliged to Dr. Sari Kujala for her assistance during the research work and for making my thesis writing as pleasant as possible.

A big thanks of course goes to the company representatives I interviewed for the purpose of this thesis.

I would like to express my sincere gratitude to my parents and my sister for their continuous support and motivation in my life, without whom this journey would have been really difficult. I take this opportunity to thank people in my life- family, friends, mentors and professors who have been part of my development process. Thanks for all the smiles and happiness.

Espoo, September 2019

Nagadivya Balasubramaniam

Contents

1 Introduction 1

1.1 Motivation 1

1.2 Research problem and questions 2

1.3 Scope of the thesis 3

1.4 Structure of the thesis 4

2 Research Methods 6

2.1 Literature study 6

2.2 Empirical study 7

2.2.1 Case companies 7

2.2.2 Research process 8

2.2.3 Data collection and analysis 9

3 Literature Study 12

3.1 Evolution of computer and information ethics 12

3.2 Ethical issues of AI 15

3.3 RE activities and quality requirements 18

3.3.1 Overview of RE activities 18

3.3.2 Quality requirements in RE 20

3.3.3 Quality requirements in case studies 25

3.3.4 Quality requirements of AI solutions 28

3.4 Summary 30

4 Empirical Study 31

4.1 Ethical guidelines of AI solutions 31

4.1.1 Company A 31

4.1.2 Company B 34

4.1.3 Company C 38

4.2 Ethical practices when developing AI solutions 41

4.2.1 Using multi-disciplinary development teams 41

4.2.2 Defining the purpose and impact of AI solutions 42

4.2.3 Communication with customers and people-oriented approach 42

4.2.4 Ethical design 42

4.3 Summary 43
5 Discussion *47*
5.1 RQ1: Ethical issues and quality requirements of AI solutions– Relationship 47
5.2 RQ2: Ethical guidelines for using AI in case companies 48
5.3 RQ3: Using ethical guidelines to define quality requirements of AI solutions 50
5.4 Limitations of the study 53
6 Conclusions *55*
References *58*
Appendix *63*

1 Introduction

1.1 Motivation

The current trend is to utilize AI in a variety of services and products such as healthcare services, transportation, banking etc. The progress of AI is making its impacts in almost all walks of human life. As these AI solutions are now taking charge of human activities, the concept of ethics in AI solutions comes under spotlight. The ethical questions and concerns that arise during AI solutions development when neglected lead to the ethical issues. The AI technology development has already witnessed the development of ethical issues relating to security, safety, privacy, transparency, integrity of the AI solutions (Rahwan, 2018). For example, ethical issues with autonomous cars are making headlines. The need of developing ethical AI solutions are recognized by many companies in Finland. Furthermore, in 2018, AI Finland set out an 'Ethics challenge' for companies in Finland to develop their AI ethical guidelines or principles.

After bidding bye to the long 'AI winter' we are now in the period where we are witnessing AI agents and AI-based real-world systems making human-like decisions (Vampley et al., 2018; Rahwan, 2018). The advancements provided by AI technologies have provided significant social benefits ranging in various fields (Rahwan, 2018). But the black box nature of AI technologies has raised several ethical questions among the stakeholders which concern safety, privacy, security, and level of transparency of AI solutions (Vampley et al., 2018; Rahwan, 2018). Nowadays there is an increase in people considering the ethical aspects of non-software goods purchase like avoiding goods that involve child labour, analysing the environmental impact of goods. A similar aspect is advocated for software organizations to consider ethics and ethical viewpoints of stakeholders when designing and applying software system in the real world (Aydemir and Dalpiaz, 2018). The question on which values are important to design a system to minimize the trade-offs of social moral values have led to the studies relating to ethics in artificial intelligence (Etzioni and Etzioni, 2016; Dignum, 2018).

The autonomous decision-making capability of AI systems has led to various ethical issues, especially the autonomous cars and robots making headlines now and then (Bonnemains et al, 2018). During the development of intelligent systems, designing the system to be accountable, responsible and transparent have gained special focus (Bonnemains et al., 2018). To develop a responsible AI, human and ethical values need to be embodied during the system design rather than considering it as a checklist that has to be met (Dignum, 2018).

In software and service development, requirements engineering (RE) is a key activity which involves discovering user needs and requirements, analysing and documenting the identified requirements. It is an essential iterative activity to maintain user requirements of the software up-to-date. In this thesis, we are discovering the importance of ethics when designing AI solutions and to identify the role of ethics in companies. The goal of this thesis is to explore how companies can solve potential ethical issues of AI solutions during requirements engineering activities.

1.2 Research problem and questions

Requirements engineering (RE) is a critical process of software engineering where needs of users and other stakeholders are discovered, analysed and represented as requirements. With AI solutions, there are increased need of identifying and analysing the ethical concerns and issues during their early phases of development. The goal of this thesis is to explore how companies can solve potential ethical issues of AI solutions during requirements engineering activities. This thesis follows qualitative research methodology which explores the ethical guidelines of AI used in the companies through interviews and analyse its connection to RE activities. The study is done in three case companies of different application domain that are developing AI solutions. The study aims to answer the following research problem and questions.

The research problem of this study is:

How can companies solve potential ethical issues of AI solutions during RE activities?

The research problem is divided into three research questions:

RQ1: What is the relationship between ethical issues and quality requirements of AI solutions?

RQ2: What kind of ethical guidelines do the case companies have for designing AI solutions?

RQ3: How can companies use ethical guidelines to define critical quality requirements of AI solutions?

The aim of the first question is to identify the importance of ethics by investigating ethical issues and impact of ethical issues. There are different ethical issues of AI solutions which are been reported in scientific articles, books, newspapers, blogs etc. In this part of study, what is the role of ethics and the different ethical issues are investigated and documented. This acts as a foundation for the further analysis of exploring the role of ethics in the case companies.

The second question assists in understanding the viewpoint of the case companies about ethics during AI solution development. Also, the goal of the empirical study is to explores the ethical guidelines developed by the case companies for designing AI solutions. The purpose of this empirical study is to analyse which potential ethical issues of AI the case companies have focused on, and what kind of ethical guidelines they have defined for solving these issues during AI solution development.

The final question focuses on identifying the link between the ethical guidelines of the case companies and quality requirements. As the initial study revealed that most of the ethical issues with AI solutions are related to the quality requirements like security, privacy and safety. Through this question, we formulate the potential link in using the ethical guidelines for defining the critical quality requirements of AI solutions.

1.3 Scope of the thesis

The thesis focuses on how the ethical viewpoint and ethical guidelines devised by case companies can be utilized in the RE process of AI solution design and development. The ethical issues have to be identified and solved during the initial stage of solution design to avoid major changes in the later stages. The different ethical issues are identified through the literature studies. The recently formulated ethical guidelines by the case companies are explored via empirical studies. The ethical guidelines of AI and the viewpoint of the case companies of using the ethical guidelines in their projects are identified and documented in this thesis.

The results of the thesis will provide ways of linking ethical guidelines of AI solutions with RE activities of solution development. This thesis does not update the existing RE process. The thesis aims to explore how companies can solve potential ethical issues of AI solutions during requirements engineering activities. Table 1 gives the illustration on scope of the thesis relating to the three research questions of this thesis.

Table 1: Scope of this thesis

Research questions	Literature Study	Empirical Study
RQ1	**X** (Studies showing different ethical issues arisen of AI and how ethics evolved with technology advancements)	
RQ2		**X** (Interviews with company professionals about their companies' AI ethics standpoint)
RQ3	**X** (Studies highlighting the RE activities and quality requirements during RE phase of software design)	**X** (Interview results showing the ethical guidelines of AI and practices in case companies)

1.4 Structure of the thesis

In Section 1, the research problem and the research questions are introduced. This is one of the important sections as the entire study is built around the research problem and the questions.

Section 2 explains the methods used to conduct the research in this thesis. First, the approach used for literature review and the sources used to retrieve relevant information are summarised. Next, the qualitative research methods used in the empirical study are elaborated. The case study approach and the description about the companies are accounted in this section.

Figure 1 gives an overview of how the structure of thesis is related to the research question of this thesis. Section 3 describes the results of literature study. Section 3.1 focuses on the evolution of computer and information ethics and in figure 1, it is related to all the three research questions. The basic knowledge on computer ethics is indispensable to ensure clarity during the discussion. The ethical issues of AI mentioned in the scientific articles are identified and categorised in Section 3.2. The third section of the literature review gives a holistic explanation on different RE activities and quality requirements. Finally, the summary section (Section 3.4) delineates the answer for the RQ1 by showing the interrelation between ethical issues and quality requirements of AI solutions.

Section 4 examines the empirical study done in relation to the RQ2 and RQ3. The empirical study of this thesis utilised multiple-case study approach. Section 4.1 summarises the ethical guidelines developed in the case companies. The overview of ethical guidelines and ethical practices sections (Section 4.2 and 4.3 respectively) discusses on the ethical guidelines that are emphasized by all the three companies. The ethical practices that complement the ethical guidelines are reported. The identification of ethical guidelines and ethical practices in the case companies complies with the RQ2. The section concludes by summarizing the lesson learned from the analysis of the different case company's viewpoint on ethics and its relation to the quality requirements (RQ3).

Section 5 holds the discussion on how the research questions are answered based on the literature and empirical study of this thesis. The summarization on identifying the relationship between AI ethics and RE is presented in this section. The last section of this thesis mentions about the limitations of the study. Section 6 include conclusions of this thesis. The scope for further research is also highlighted as a section of this chapter.

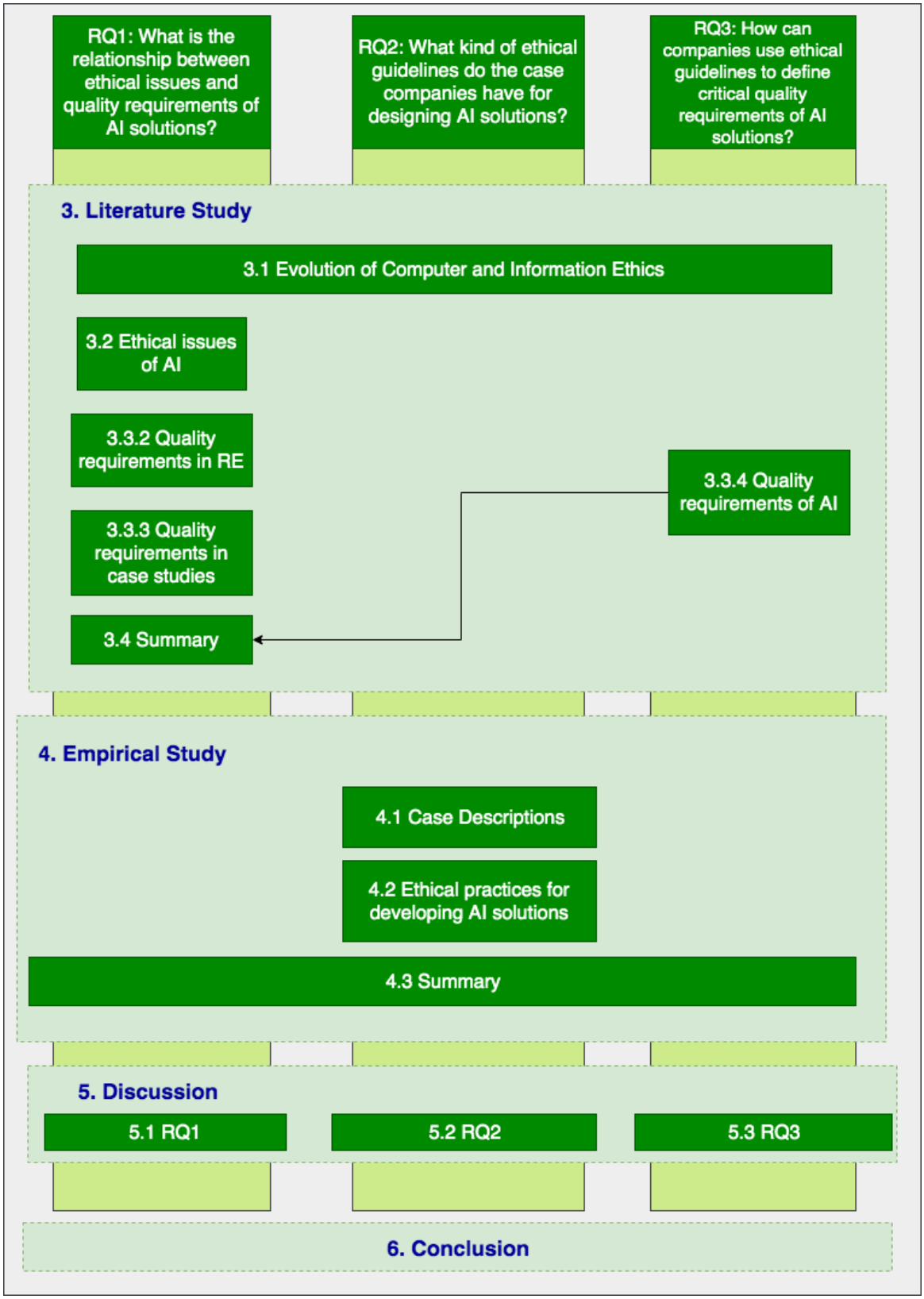

Figure 1: *Relationship between research questions and structure of the thesis*

2 Research Methods

This chapter summarizes the research approach utilized in this thesis. Firstly, the literature review and the process to that is elaborated. The empirical study process of this thesis is discussed in the following section.

2.1 Literature study

The literature study of this thesis focuses on: First, analysing why it is important to consider ethics when designing AI solutions. Next, the study progresses by delineating RE activities and different quality requirements when designing software solutions. The literature review part of the thesis concerns in answering the research questions 1 and 3 (RQ1 and RQ3). The research relied predominantly on the scientific articles, books and information from official site of companies. As AI ethics and digital ethics are relatively new topic, there was limitations in finding related scientific publication on the topic. The books on computer and information ethics were used to understand the role of ethics in the field of computer science. The fundamental textbooks and articles on requirements engineering were used to get an overview of the topic.

The more in-depth information analysis of the literature review was done for answering the research questions. Few articles were selected based on the key concepts of RQ2, so that a strong theoretical understanding was gained when documenting the results. The research papers used in this study are retrieved from Aalto university database, Aalto digital library (LibGuides) and Google Scholar. The following list shows the sites used to collect the publications relevant to the thesis.

- ScienceDirect (Elsevier) https://www.sciencedirect.com/
- SpringerLink https://link.springer.com/
- IEEE Xplore https://ieeexplore.ieee.org/Xplore/home.jsp
- ACM Digital Library https://dl.acm.org/dl.cfm?coll=portal&dl=ACM
- Scopus https://www.scopus.com/search/
- Google Scholar https://scholar.google.fi/

This study focuses on two areas i.e. ethics and requirements engineering from a big picture perspective. First, the articles were searched relating ethics and requirements engineering separately. Then, articles and publications which relates to ethics and requirements engineering together were explored.

The common keywords used in the search are shown in the Table 2. They are divided into two categories based on the two study areas of this thesis. In the ethics category, keywords ‘computer ethics’ and ‘information ethics’ fetched a vast range of articles. But the keywords search on ‘AI ethics’ resulted in few articles highlighting the newness of the topic the research field.

To answer the first research question, the was divided into three steps, the first step was to understand what is ethics and the evolution of ethics in the computer field. This resulted in articles indicating the varying trend of ethics in computer field. For the next step, ethical issues relating to AI are searched which resulted in couple of high-quality scientific articles. In the last step, information relating to

quality requirements and its association with AI are searched to know the existing relationship viewpoint on these topics.

Table 2: Categories and keywords of literature review

Category	***Keywords***
Ethics	Computer ethics, information ethics, AI ethics, ethical issues with AI, ethical issues in autonomous systems, Ethics AND AI
Quality requirements	Quality requirements AND requirements engineering, quality requirements in companies, quality requirements AND AI, non-functional requirements AND AI, NFR AND autonomous systems, Quality requirements AND AI ethics

There are not many articles relating to the quality requirements and AI ethics or about having ethical perspective in requirements engineering. Hence, the related information is collected from conference papers which provide the results of their ongoing research in this field. In this research, literature can be only be used as a theoretical starting point for identifying the relationship between quality requirements and ethics (of AI).

2.2 Empirical study

This chapter introduces the case companies which are involved for the empirical study of this thesis. The research utilises the interview data analysis and document analysis which are explained in detail in this section. In addition, this section also explains how the data was collected and analysed for the empirical study.

2.2.1 Case companies

This section introduces the background of the case companies and provides an overview of the case companies. Table 3 shows the companies of the multiple-case study done in this thesis. The size of the company and their respective application domain are presented in Table 3. The research process is detailed in the following Section 2.2.2.

Table 3: Companies of the multiple-case study

Company	Number of employees	Application domain
A	~12 300	Banking
B	~22 500	Retail
C	500	Software Consultancy

Company A is one of the largest financial service providers in Finland with millions of customers. The company provide banking, insurance and retail services. The headquarters is in Helsinki, Finland and the company employs over 12 thousand people and has branches all over Finland. The company has formed data science teams around four years back and started working on ethical AI since 2017. The AI team is currently developing new kind of applications and services in financial sector. Interviewee is handling ethical analytics and ethical software design because the banking and financial services are one of the sensitive areas.

Company B is a Finnish retail company involved in car trade, food trade and building trade. The company has over 20 thousand employees in Finland and has collaboration with various organisation chains. The company endorses the open and sustainable way of working. They are specialists in supporting three different businesses of the company. The team comprises of 25 people with different skill set and capabilities. AI ethics is been in the discussion for couple of years already in the case company. The company has a special AI solution which supports their customers make sustainable choice and have healthier lifestyle.

Company C is a software consultancy company with Finnish origin. The company currently has over 500 employees working around Europe. The company designs and delivers new digital services and products. The company emphasizes the quality process for doing machine learning and AI projects. This company is providing AI ethics coaching to the employees. The data science team has around 20 people who are mainly data scientists and coders.

2.2.2 Research process

To answer the research questions, the study was carried out using qualitative methods (Cassel & Symon, 2012) to identify companies' current situation relating to ethics. We carried out a qualitative case study research adhering to the approach of Eisenhardt and Graebner (2007). We utilized Alsaqaf et al.'s (2018) empirical research method to describe our research process. Figure 2 gives an overview of the research process used in the empirical study.

In order to explore the RQ2 to discover '*what kind of ethical guidelines that the case companies have for designing AI solutions*', interview study is carried out in the case companies. The interview data and documents about ethical guidelines in the companies collected during the interview are analysed.

The analysis of ethical guidelines in the case companies helped in summarizing the ethical guidelines of AI solutions. Section 2.2.3 explains the data collection and analysis process in detail.

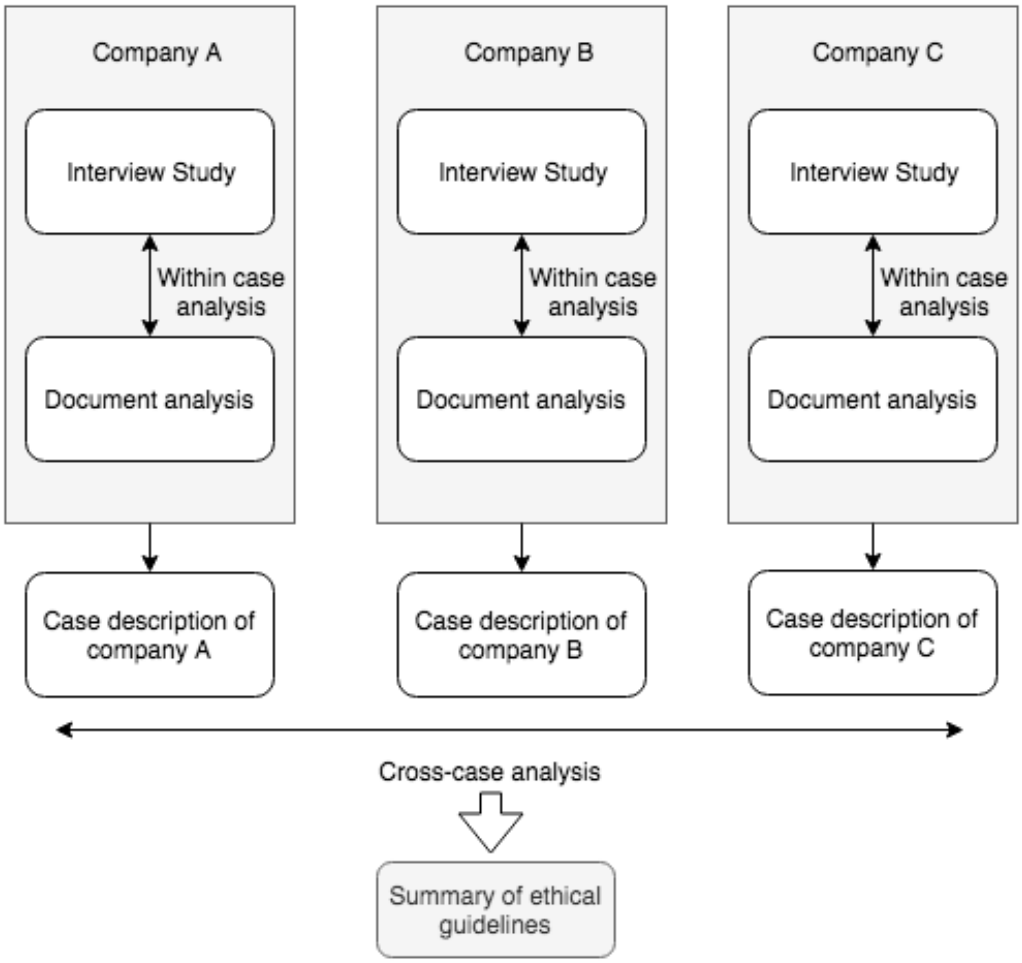

***Figure 2**: Overview of the research process*

2.2.3 Data collection and analysis

For the purpose of data collection, semi-structured open-ended in-depth interviews are planned based on the guidelines of Boyce and Neale (2006). As a first step, we made an interview plan describing: 1) the objective of the interview, 2) the kind of potential interviewees who could be able to elaborate more on the topic from company perspective, and 3) the set of questions which could help us in collecting information we intend.

The unit of analysis in our empirical research process was company-level. A senior researcher accompanied as an interviewer along with the author of this thesis during the interview sessions. As we wanted to understand the role of ethics during software development in different domains, we decided to include interviewees representing different application domain. As per Yin (2013), Boyce and Neale (2006), a multiple-case study could help in exploring in-depth knowledge on the relatively new issue. As Yin (2013) suggests, the number of cases influence the research, the three cases of

different domains were chosen for this thesis to get a big picture of the current situation of ethical guidelines in the case companies.

Secondly, interview protocol was devised with the procedure of collecting related documents after each interview. The interview questions were designed based on the research goal of the study and they were validated by three senior researchers. The interview questions were improved and finalized after iterative feedbacks from the senior researchers. Furthermore, a pilot testing of interview questions with an expert was carried out to check the feasibility of questions. There were no changes made to the interview questions after the pilot testing stage. The pilot testing interview is not included in this multiple-case study. The interview questions were organised in two parts. The first part was about the organizational context of the interviewees, and the second part focused on ethics and ethical guidelines.

Third, for the purpose of data collection, we interviewed three interviewees in total from three different organizations. The interviewees had a high-level view on AI related ethics in their respective companies and were able to describe the big picture of ethical guidelines in their companies. The interviewees also had close collaboration with professionals of different roles such as designers, data scientists, and developers. An overview of the case description elaborated in the previous section is summarized in the Table 3.

The interviews were conducted in late 2018 and early 2019. The length of the interviews varied from 60 to 80 minutes. Each interview began with a short self-introduction of interviewers followed by introducing the research objective and goal of the interview study. We also had an agreement to audio-record the interview and assured anonymization in results. After the interview, interviewees shared the set of public and internal documents of their company (Table 4).

Table 4: Multiple-case study companies and their related documents information

Company	Name of document	Availability
A	Ethical guidelines for the use of AI	Public
B	Ethical principles for utilizing AI	Public
C	AI Ethics training	Internal
	Data ethics canvas	Public

Next step was the analysis of the data we collected from interviews and documents. The audio files from the interviews were transcribed and notes-taken during the interview was attached to transcribed document. The data analysis process of this study followed Charmaz's (2006) grounded theory method. Charmaz's (2006) method was chosen for its suitability of our qualitative research. The qualitative open coding process (Charmaz, 2006) was employed to make the analytical interpretation of the collected data. First, the initial segments coding was done for texts in each transcript. In the next stage, the descriptive codes from initial coding were grouped together into high-level categories. In addition, we coded the information from documents separately as well. The categorization of codes

was discussed between the researchers to avoid any possible ambiguities. Charmaz's (2006) grounded theory method was used for analysis by open coding of the qualitative interview data.

As there were two different sources of data (interviews and documents) from each company, the Eisenhardt (1989) research method was applied in the data analysis of this study. The within case analysis was performed with the codes categorized from the interview transcripts and the documents. The case description report was written for each company based on the results from the within case analysis. Thereafter, based on the case description written for each company, the cross-case analysis method was employed. During cross-case analysis, evidence data from one case description was compared with other cases. Figure 2 gives an overview of the research process employed in this study.

When documenting the results, the ethical guidelines specified in the companies are delineated to the reported ethical issues by adhering to the Conklin's Dialog Mapping technique for the process of qualitative data structuring (Alsaqaf et al., 2018; Conklin, 2003). This technique assisted us in organizing the results of our data analysis by forming the ethical issues-ethical guidelines-arguments structure (Alsaqaf et al., 2018). The ethical issues and ethical guidelines are those derived from our data analysis, and the arguments are from interviewees thoughts and comments on the respective ethical guidelines.

3 Literature Study

This chapter is divided into three sections, the first section 3.1 describes the evolution of computer and information ethics. The second section 3.2 gives an overview of ethical issues of ethical issues with AI solutions and the third section 3.3 describes the RE activities and quality requirements during software systems design.

3.1 Evolution of computer and information ethics

Figure 3 represents the timeline of evolution in computer and information ethics. Computer and information ethics are expanding with new applied ethics theories like digital ethics, ethics in artificial agents, AI ethics which are stemming from their roots as a result of the emergence of new technological advancements (Allen and Roberts, 2010). This continuous growth of AI applications can have significant impact on the society which demands ethical check in the early stages of their design and development.

In the early 1940s, Norbert Weiner during World War II introduced a new research theme called cybernetics which emerged as a science of information feedback systems (Iqbal and Beigh, 2017; Stanford Encyclopedia of Philosophy, 2001). Weiner remarked machines, humans, and animals as cybernetic entities, in which machines included digital computers and robots (Spinello, 2012). He developed various ethical insights on technology which we now term as information and communication technology (ICT) (Stanford Encyclopedia of Philosophy, 2001). His ideas were way ahead of time and were ignored for almost a decade. In 1950s Weiner's books on cybernetics served as a foundation for new computer related ethics researches. Around mid-1960s Donn Parker in his books enumerated the various computer crimes (Iqbal and Beigh, 2017). The need for regulatory measures, laws and guidelines to handle those crimes and unethical occurrences were raised in this period (Iqbal and Beigh, 2017; Stanford Encyclopedia of Philosophy, 2001).

In mid-1970s, Walter Maner coined the term computer ethics and defined it as a new field of study as the issues raised around computers were persistent (Iqbal and Beigh, 2017; Spinello, 2012). In mid-1980s James Moor's path changing article 'What is computer ethics?' broadened the field of study by incorporating 'computer technology' which included software, hardware, networks, etc. Moor included consideration of human values with ethical issues relating to computers (Iqbal and Beigh, 2017). In 1988, Robert Hauptman used the term 'information ethics' for the first time (Iqbal and Beigh, 2017). The study then focused on creation, organization and use of information about human conduct in the society. In the 1990s, Donald Gotterbarn developed a new perspective of connecting professional ethics with computer ethics. The exemplary work of him with ACM and IEEE progressed in the publication of 'Code of Ethics and Professional Conduct' by ACM in 1992 (Iqbal and Beigh, 2017).

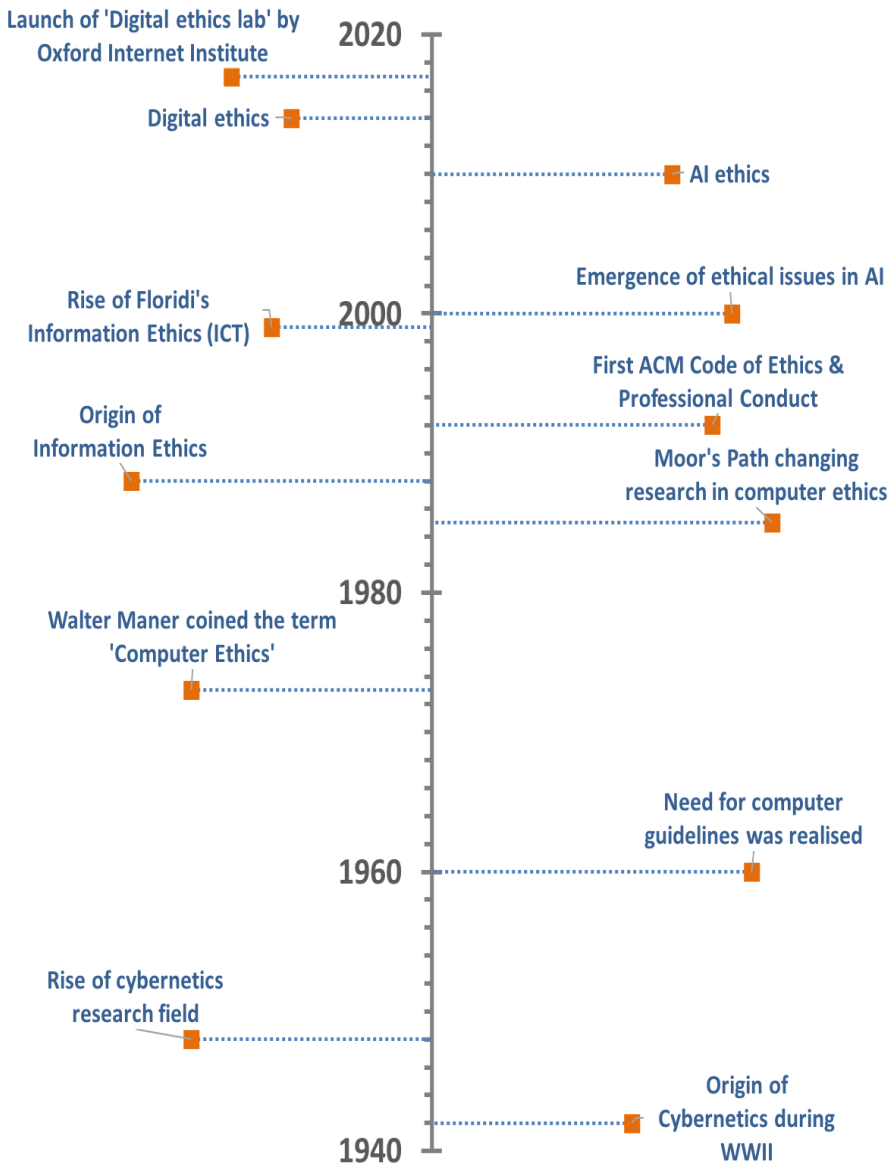

Figure 3: *Timeline of evolution of computer and information ethics*

After the information revolution, the ICT field developed rapidly with various inventions (Floridi, 2001). As the inventions like telephone, camera, computers, internet, email, broadband, wireless networks, etc. have become a part of our everyday life, various issues surfaced with these ICT developments. Spams, challenges in censorship, privacy issues, consumer backlash, hacking, malware, cybercrimes, and cyber-attacks are some ethical issues relating to these ICT inventions (Quinn, 2014). Computer and information ethics were adopted during the development of new technologies to identify the potential issues and to curb those risks (Quinn, 2014).

Luciano Floridi's 'information ethics' which emerged as new computer ethics theory occurred in the late 1990s and the early 2000s (Iqbal and Beigh, 2017). Floridi's information ethics is spun around various traditional ethics theories and aimed to avoid conflicts between those theories (Spinello, 2012).

Information ethics groups issues on the multiplicity of occurrence based on the interpretation of information present in the environment (infosphere) (Iqbal and Beigh, 2017). It has to be emphasized that from late 1990s the term 'computer ethics' and 'information ethics' are almost used interchangeably or referred as 'Computer and Information ethics' though they were two separate theories in the past (Spinello, 2012). The early 2000s witnessed the rise of ethics issues in human-computer interaction and artificial intelligence.

In the early 2010s Bostrom and Yudkowsky (2011) published their study 'Ethics of Artificial Intelligence' concentrating on the ethical considerations relating to the autonomous or thinking machines. The need for safety of these AI technologies and moral status of these intelligent machines are highlighted. They also indicated the requisite of new design requirements like transparency and predictability (Bostrom and Yudkowsky, 2011). From mid-2010s, digital ethics is one of the indispensable topics and the shift of ICT from business supplementary to business enabler portrays the development of ICT technologies. The ethical debate raised in the society as the consequence of intentional or unintentional misuse of technologies is the driving factor to adopt digital ethics (Shetty, 2017). In 2017, Digital Ethics Lab led by Luciano Floridi was launched by Oxford Internet Institute. It concentrated on analysis focusing on bringing positive impact with digital innovations and eliminate risks of it from the society (Oxford internet institute).

3.2 Ethical issues of AI

This decade is facing a rapid advancement in AI and machine learning, and digitalization making a robust growth, all of which led to various technology-oriented ethical issues. There have been ethical issues appearing when developing and deploying the AI solutions and regarding their regulatory mechanisms (Rahwan, 2018).

All the technological advancements impacted by easing and facilitating our daily life, and opened the market for various opportunities. They also disturbed the moral boundary with its ability to make decisions and results on its own (Royakkers et al., 2018). People and their information are monitored and used more than ever before through various smart devices and applications used in their daily life routine. Figure 4 is the representation of different ethical issues with the AI solutions and trust being the influential factor of all the ethical issues.

Figure 4: *Ethical issues of AI solutions*

Privacy

When developing AI solutions, *privacy* sprouts in as a key area of ethical issues. The collection of volumes of personal data from different sources is the cornerstone for building many AI solutions. Any potential misuse of data could lead to major privacy threats (Jones et al., 2015). Privacy issue relating to AI is bound to produce both individual and societal level impact, for example, digital home issues like personal data collection of users with manipulated consent, digital spying of citizens, even

in simple privacy settings in social media applications, etc. (Etzioni & Etzioni, 2016; Royakkers et al., 2018).

Autonomy

Though branched out from privacy, *autonomy* is considered as a distinct ethical issue because of the number of ethical issues surrounding them. Autonomy in AI context refers to the ability of the computers to make real-time decisions based on the environmental inputs without human involvement (Etzioni & Etzioni, 2016). Some notable ethical concerns relating autonomy are: 1) the extent that technologies can influence humans (Royakkers et al., 2018), 2) level of consideration of personal autonomy e.g. surveillance of workers (Palm, 2009), 3) capacity of individuals to make their own choice (Jones et al., 2015), 4) possibilities for man out-of-loop and its impact (Royakkers et al., 2018).

Anonymity

The ethical issue in *anonymity* (Doyle & Veranas, 2014) relates to the public value of being anonymous which also relates to ethical issues privacy and autonomy. It is argued that anonymity helps in achieving autonomy (Doyle & Veranas, 2014). The digital services and surveillance have put anonymity in stake. For example, CCTV when associated with face recognition technology, which captures video footage of people and matches with data and images from other sources. This is a direct threat on the anonymity of individuals (Doyle & Veranas, 2014).

Transparency

Transparency in data used for AI decision-making is proposed as a solution to protect individual's autonomy and privacy. From an ethical perspective, by adopting transparency in AI solutions it should inform both positive and negative impact on their potential users (Jones et al., 2015). It is also advocated that transparency unlades manipulation and shows why an AI solution make certain decisions or provide certain choices (Royakkers et al., 2018). In ICT, transparency represents the significance of the stakeholders' right to know. And the 'right to know' leads the transparency and trust go hand in hand (Elia, 2009).

Security

The crimes occurring with the digitalization and AI trend are imposing series of *security* challenges. The security challenges are seen to coexist between privacy issues in examples. Focusing on security requirements during AI solutions design reduces privacy related issues (Jones et al., 2015). With IoT utilization in boom, issues relating to security are growing complex. For example, hacking a coffee machine at home can help the hackers to open the front door (Royakkers et al., 2018). Pieters (2011) indicates his insights on how explanation of security measures can enhance the information security and trust in AI solutions. Identification and inclusion of essential security measures should be a key ethical responsibility for any AI solution to gain users' trust (Jones et al., 2015).

Safety

Safety in AI has triggered various ethical and societal issues. When driving autonomous cars, developing computer games with VR, employing GPS technologies to track elderly patients, and in other social and daily life services physical and psychological safety of users is crucial (Jones et al., 2015; Royakkers et al., 2018; Pieters, 2011; Bostrom & Yudkowsky, 2011). It is important to prioritize users' safety at any given point of time during the development and deployment of the AI solutions (Pieters, 2011). Pieters (2011) and Jones et al. (2015) highlighted that the safety, security and privacy needs have to be balanced because any trade-offs on these turns to an unpleasant example.

Justice

Jones et al. (2015) defines *justice* in intelligent environment as "the moral obligation to act on a fair adjunction between conflicting claims". The technology should cater everyone without any discrimination of gender, age, accessibility etc. (Jones et al., 2015). Nowadays, with the data collected through biometrics and other sources, stigmatization of individuals is done without the person being aware of it. Many AI solutions are well-known for its wrong judgments which happened because of wrong data exclusion or discrimination (Royakkers et al., 2018). Another unjust act relating to the AI solutions are profiling of users based on their data. In profiling examples, it is always certain that particular user group is disadvantaged and their future options are heavily influenced (Royakkers et al., 2018). Justice as an ethical issue can be fixed by enforcing equality on every user groups (Jones et al., 2015).

Dignity

Dignity as an ethical issue in AI solutions refers to human dignity that is being compromised or not considered at all. This trend of robotics and AI possess a potential for dehumanization and desocialization of people. For example, a caregiver robot could tamper the human values of care and empathy. With AI solutions comes the risk of ultimate behavioral change of people by discarding their moral values. Like, standard awareness of individuals is reduced with a driver support system warning driver to fasten their seatbelts (Royakkers et al., 2018). Undermining human dignity could relate to other areas of ethical issues like privacy, justice and equality (Jones et al., 2015; Pieters, 2011).

Trust

All these ethical issues have a foremost connection with *trust* which serves as a core concept. Trust is a form of assurance given to users which strengthens the reliability by eradicating the risks of using an AI solution (Pieters, 2011). The lack of visibility or simply the black box nature of these AI solutions leaves the users bewildered on certain suggestions made by AI solutions. Explanation is seen as the key idea to build trustworthy AI solutions (Pieters, 2011; Elia, 2009). In some AI solutions to attain security and privacy, explanation is coupled with transparency to showcase how and why a decision or suggestion is made to the users (Elia, 2009). When deploying AI solutions trust from users is achieved by incorporating adequate measures in ethical issues relating to security, privacy, autonomy, and transparency (Jones et al., 2015).

3.3 RE activities and quality requirements

3.3.1 Overview of RE activities

What is RE?

The definitions on requirements engineering has been transforming over period. In 1997, Zave defines the requirements engineering (RE) as branch of software engineering that focuses on the real-world goals, application and constraints of software systems. The other definition is that requirements engineering is the activity of "discovering, documenting, and maintaining a set of requirements" for a computer-based system (Sommerville and Sawyer, 2003).

Nuseibeh and Easterbrook (2000) mention that RE is a process of identifying the needs of stakeholders and documenting it. This documentation is analysed, communicated and then the requirements are implemented. The current requirements engineering has its emphasize on customer value creation where the different 'elements of values' are identified by analysing user or customer processes (Almquist et al, 2016). Requirements engineering from being a discipline that concentrated on specifications and requirements is evolved into important foundation block of the software development process (Paech and Kerkow, 2004).

Requirements

Requirements of a software and service can be defined in many ways. Requirements signify the attributes or properties on, how the system should perform (Sommerville and Sawyer, 2003). Requirements are user needs that the software or system should fulfill and these needs are defined through context analysis, functional specification and design constraints methods and criteria (Ross and Schoman, 1977). The two kinds of requirements are: Functional requirements and Non-functional requirements (NFRs).

Functional requirements describe the intended behaviour of the system. It specifies what the system should do (Sommerville and Sawyer, 2003). The behaviours are specified as tasks or features, services like UI requirements, database requirements. Collection and analysis of functional requirements are focused in the initial stages of software or system development.

Non-functional requirements are specified as constraints on how the functional requirements are implemented (Sommerville and Sawyer, 2003). These requirements are helpful in increasing the software quality which indicate its relevance during solution development. They can be classified to broad range of categories which are domain specific like security, performance, usability, availability, portability, interoperability etc. An unfortunate part about NFRs are they are discovered clearly only on the later stages of development cycle because of the unclassified documentation during the requirements gathering phase (Huang et al, 2006). The non-functional requirements are also called as quality requirements or quality attributes.

RE activities

The requirements engineering processes include set of activities which are carried out iteratively to satisfy the customer needs. Sommerville and Sawyer (2003) specifies the following key RE activities:

- ***Requirements elicitation*** (Sommerville and Sawyer, 2003; Paetsch et al., 2003; Loucopoulos and Karakostas, 1995; Kotonya and Sommerville, 1998; Svensson, 2011)

  The RE process models of Loucopoulos and Karakostas (1995), Kotonya and Sommerville (1998) and Sommerville and Sawyer (2003) mention requirements elicitation as the starting RE activity. Requirements are discovered in this stage. The requirements are gathered from key stakeholders who may be anyone who will be impacted by the system e.g. customers, users, third party regulators (Svennson, 2011). Starting with an appropriate knowledge on domain aids in gathering the requirements. The different sources to gain domain knowledge are existing software system, similar existing applications, literature, national standards etc. (Loucopoulos and Karakostas). The techniques used for eliciting the requirements from stakeholders are interviews, ethnography, focus groups, brainstorming, prototyping (Paetsch et al., 2003; Svennson, 2011).

- ***Requirements analysis*** (Paetsch et al., 2003; Loucopoulos and Karakostas, 1995; Kotonya and Sommerville, 1998; Sommerville and Sawyer, 2003)

  Loucopoulos and Karakostas (1995) defines requirements analysis as:

  > *"requirements analysis is the reasoning process which attempts to understand the requirements of a problem domain in order to synthesise a solution for a system which will satisfy the needs of its users"*

  In this RE activity the requirements are analysed with stakeholders or users for feasibility and the essential needs are identified. During the analysis, conflicting requirements and inconsistent requirements are also walked through (Sommerville and Sawyer, 2003). Key techniques used in requirements analysis are joint application development (JAD) sessions, requirements prioritizing and modelling (Paetsch et al., 2003).

- ***Requirements specification, representation or documentation*** (Kotonya and Sommerville, 1998; Sommerville and Sawyer, 2003; Svensson, 2011)

  This activity is referred by different terminologies in RE process models which all relates in performing the same task. Svensson (2011) defines requirements specification as an activity that includes "*documenting the elicited functional and non-functional requirements in detail"*. The user needs collected are modelled and documented after eliciting and analysing the requirements (Kotonya and Sommerville, 1998; Sommerville and Sawyer, 2003). The requirements can be documented in different ways like use cases, requirements model, formal specifications etc. (Svensson, 2011).

- ***Requirements validation*** (Loucopoulos and Karakostas, 1995; Sommerville and Sawyer, 2003; Sommerville, 2010; Svensson, 2011)

  Requirements validation is classified as the final stage of RE process that involves verifying whether the requirements are consistent and complete and are in line with the user needs (Sommerville and Sawyer, 2003; Svensson, 2011). Loucopoulos and Karakostas (1995) classifies two set of activities for requirements validation. They are: preparing the setting for experiment and analysing the results from the experiments. Any problems with the requirements are identified in this stage and there are different techniques to validate them. Svensson (2011) mentions the use of reviews and inspections, and employing checklists in order to identify the defects, missing parts, inconsistencies in the requirements and to improve the quality.

- ***Requirements management*** (Sommerville, 2010; Svensson, 2011)

  Svennson (2011) defines requirements management as:

  "Requirements management is the process of understanding, reviewing, analysing, controlling, and communicating to all stakeholders' changes to a system's requirements"

  This is an ongoing process of understanding and controlling changes to the system requirements (Sommerville, 2010). The system requirements change or new requirements emerges during the software system development is inevitable. The change request on old requirement or the new requirements are analysed and then approved as system requirement. Requirements traceability is also the part of requirements management which keep tracks in maintaining the link between dependent requirements (Sommerville, 2010; Svensson, 2011).

In addition, Nuseibeh and Easterbrook (2000) recommend activities like communicating requirements, agreeing requirements and evolving requirements as core RE activities. The selection of activities during RE processes depends on the application domain and people working on RE in organisations (Sommerville and Sawyer, 2003). The RE processes occurs with different stakeholders, designers, users, project managers and other potential contributors. This activity helps in deciding what has to be done and decides on tools to carry out processes (Aurum and Wohlin, 2003).

3.3.2 Quality requirements in RE

What is quality requirement?

The different types of requirements are usually classified as functional requirements or non-functional requirements. Non-functional requirements (NFR) are also called as quality requirements (Glinz, 2007; Sommerville, 2010). The quality requirements focus on the question *'How good the software performs?'* versus the functional requirements question on *'What the software does?'* (Paech &Kerkow, 2004). There are multiple definitions for quality requirements and the Table 5 provides an overview of quality requirements definitions which are selected from Glinz's (2007) extensive study on quality requirements (Svensson, 2011).

Table 5: Definitions of quality requirements

Reference	Definition
Sommerville (2010)	Quality requirements are not directly concerned with specific functions delivered by the system. They may relate to emergent system properties such as reliability and response time. Alternatively, they define constraints on the system such as capability of I/O devices and the data representation used in system interfaces
Thayer and Dorfman (1990)	In software engineering, a software requirement that describes not what the software will do, but how the software will do it, for example, software performance requirements
Mylopoulus et al. (1992)	"… global requirements on its development or operational cost, performance, reliability, maintainability, portability, robustness, and the like. (…) There is not a formal definition or a complete list of non-functional requirements."
Kotonya and Sommerville (1998)	Requirements which are not specifically concerned with the functionality of a system. They place restrictions on the product being developed and the development process, and they specify external constraints
Lauesan (2002)	Quality requirements specify how well the system performs its intended function

Glinz (2007) definition on quality requirements after synthesising 13 selected definitions on quality requirements is:

"A non-function requirement is an attribute of or constrain on a system".

There are several standards published in order to classify the quality requirements. For example, ISO/IEC 9126, McCall, FURPS, and the like. The ISO/IEC 9126 model signifies 6 main characteristics and 27 subordinate characteristics. Table 6 represents the ISO/IEC model:

Table 6: ISO/IEC 9126 model of quality requirements classification

Characteristics	***Sub-characteristics***
Functionality	Suitability, accuracy, interoperability, security, functionality compliance
Reliability	Maturity, fault tolerance, recoverability, reliability compliance
Efficiency	Understandability, learnability, operability, attractiveness, usability compliance
Maintainability	Analysability, changeability, stability, testability, maintainability compliance
Portability	Adaptability, installability, replaceability, coexistence, portability compliance

For example, Chung et al. (2009) highlights a framework that focuses on goals of the system in quality requirements perspective. The categories of QRs are:

1. Interface requirements - User interface and its qualities i.e. User-friendliness
2. Performance - Time space bounds, reliability, security, survivability
3. Operating requirements- Physical constraints, system accessibility and maintenance
4. Lifecycle requirements - Quality of design, limits on development
5. Economic and political requirements.

FURPS model developed by Hewlett-Packard classifies the non-functional requirements or quality attributes with root categories as functionality, usability, reliability, performance and supportability with each having set of branch categories (Chung et al., 2009). Every company adopts some model or scheme based on their domain and system to evaluate the software quality attributes. Apart from this, Mairiza (2010) detailed study on quality requirements reported 114 quality requirements that are referred in various literature. The analysis of identified requirements resulted in segregating the most commonly used NFRs. The Table 7 shows the top 5 commonly used NFRs, its definition and their relating quality attributes and Figure 5 shows the list of 114 QRs collected from various sources by Mairiza et al. (2010).

Table 7: The most commonly considered NFRs (Mairiza et al., 2010)

NFRs	Definition	Attributes
Performance	requirements that specify the capability of software product to provide appropriate performance relative to the amount of resources needed to perform full functionality under stated conditions	response time, space, capacity, latency, throughput, computation, execution speed, transit delay, workload, resource utilization, memory usage, accuracy, efficiency compliance, modes, delay, miss rates, data loss, concurrent transaction processing
Reliability	requirements that specify the capability of software product to operates without failure and maintains a specified level of performance when used under specified normal conditions during a given time period	completeness, accuracy, consistency, availability, integrity, correctness, maturity, fault tolerance, recoverability, reliability, compliance, failure rate/critical failure
Usability	requirements that specify the end-user-interactions with the system and the effort required to learn, operate, prepare input, and interpret the output of the system	learnability, understandability, operability, attractiveness, usability compliance, ease of use, human engineering, user friendliness, memorability, efficiency, user productivity, usefulness, likeability, user reaction time
Security	requirements that concern about preventing unauthorized access to the system, programs, and data	confidentiality, integrity, availability, access control, authentication
Maintainability	requirements that describe the capability of the software product to be modified that may include correcting a defect or make an improvement or change in the software	testability, understandability, modifiability, analyzability, changeability, stability, maintainability compliance

1. Accessibility/Access Control
2. Accountability
3. Accuracy
4. Adaptability
5. Additivity
6. Adjustability
7. Affordability
8. Agility
9. Analyzability
10. Anonymity
11. Atomicity
12. Attractiveness
13. Auditability
14. Augmentability
15. Availability
16. Certainty
17. Changeability
18. Communicativeness
19. Compatibility
20. Completeness
21. Complexity/Interacting Complexity
22. Composability
23. Comprehensibility
24. Comprehensiveness
25. Conciseness
26. Confidentiality
27. Configurability
28. Conformance
29. Consistency
30. Controllability
31. Correctness
32. Customizability
33. Debuggability
34. Decomposability
35. Defensibility
36. Demonstrability
37. Dependability
38. Distributivity
39. Durability
40. Effectiveness
41. Efficiency/Device Efficiency
42. Enhanceability
43. Evolvability
44. Expandability
45. Expressiveness
46. Extendability
47. Extensibility
48. Fault/Failure Tolerance
49. Feasibility
50. Flexibility
51. Formality
52. Functionality
53. Generality
54. Immunity
55. Installability
56. Integratability
57. Integrity
58. Interoperability
59. Learnability
60. Legibility
61. Likeability
62. Localizability
63. Maintainability
64. Manageability
65. Maturity
66. Measurability
67. Mobility
68. Modifiability
69. Nomadicity
70. Observability
71. Operability
72. Performance/Efficiency/ Time or Space Bounds
73. Portability
74. Predictability
75. Privacy
76. Provability
77. Quality of Service
78. Readability
79. Reconfigurability
80. Recoverability
81. Reliability
82. Repeatability
83. Replaceability
84. Replicability
85. Reusability
86. Robustness
87. Safety
88. Scalability
89. Security/Control and Security
90. Self-Descriptiveness
91. Simplicity
92. Stability
93. Standardizability/ Standardization/Standard
94. Structuredness
95. Suitability
96. Supportability
97. Survivability
98. Susceptibility
99. Sustainability
100. Tailorability
101. Testability
102. Traceability
103. Trainability
104. Transferability
105. Trustability
106. Understandability
107. Uniformity
108. Usability
109. Variability
110. Verifiability
111. Versatility
112. Viability
113. Visibility
114. Wrappability

Figure 5: *The list of NFRs identified by Mairiza et al. (2010)*

Significance of quality requirements

Quality requirements are more critical when developing software solutions (Mairiza et al. 2010). Though the importance is always placed on the functional requirements, the defined function or feature becomes useless without the quality characteristics (Chung et al., 2009). When the quality requirements are neglected or compromised it could lead to following shortcomings (Svensson, 2011):

- The system or product may be difficult to use or can be expensive to maintain (Ebert, 1998).
- Customer satisfaction is compromised because of the low software quality.
- It increases the time-to-market and influences the cost of the software.
- It even leads to failure in the software development.

There are number of examples in history relating to the impact caused because of poorly handles QR (Svensson, 2011). The key importance of QR are:

- QR impacts the market success – When the customer requirements are satisfied the product company and customers are positively impacted (Svensson, 2011).
- QRs have direct impact on functional requirements i.e. dynamic interdependencies between requirements (Karlsson et al., 2007).

Quality requirements should be defined in the early stages of development and it is one of the critical tasks essential for the success of the developed software product (Svensson, 2011).

Challenges related to quality requirements

The challenges relating to quality requirements are found in literature (Cysneiros et al., 2004; Svensson et al., 2010; Karlsson et al., 2007; Svensson, 2011). Some of the challenges of quality requirements are:

- Gathering QR is a difficult task – QR elicitation practices are vaguely defined and are highly influenced by the functional requirements and software architecture (Svensson et al., 2010; Cysneiros et al., 2004).
- Specifying QR in natural language is challenging because communication between the stakeholders need a natural vocabulary (Cysneiros et al., 2004).
- Influence of QR interdependencies – The impact of one QR on functional requirement or other QR should be identified and also analysis on how to deal in case of issues should be considered (Karlsson et al., 2007).
- QRs influence the release planning – QR interdependencies have to be identified earlier because it affects the architectural aspects of the system (Karlsson et al., 2007).

3.3.3 Quality requirements in case studies

Table 8 signifies the quality requirements that are emphasized in the recent literature. The articles selected are case studies which are published in last 10 years. The case studies articles are selected to highlight the quality requirements that companies focus on during software development.

Table 8: Quality requirements highlighted in the selected six case study articles

Source / QRs	Svensson et al. (2009)	Poort et al. (2012)	Svensson et al. (2012)	Philips et al. (2012)	Mahmoud & Williams (2016)	Ameller et al. (2016)
Accuracy	✓			✓	✓	
Interoperability	✓		✓		✓	✓
Maintainability	✓	✓	✓	✓		✓
Performance	✓	✓	✓		✓	✓
Privacy	✓		✓		✓	
Reliability	✓	✓	✓	✓	✓	
Reusability	✓		✓			✓
Safety	✓		✓		✓	
Security	✓	✓		✓	✓	✓
Usability	✓	✓	✓	✓		✓

Accuracy

The requirements which occurred in limited case studies are accuracy, interoperability and reusability. As mentioned earlier, **accuracy** co-occur with other quality requirements like performance, reliability, safety and security (Mairiza et al., 2010; Philips et al., 2012; Svensson et al., 2012). It is more relevant in application domains like banking and finance, government and military, insurance, and transport (Mairiza et al., 2010). It is the fourth important QR from the perspective of project leaders reported in the Svensson et al.'s (2009) case study. The companies of the case study of Philips et al.'s (2012) perceived accuracy as the most important quality requirement.

Interoperability and Reusability

Interoperability is a quality requirement that is identified more relevant to the application domains like education, and insurance (Mairiza et al., 2010). Interoperability requirements are applied to manage the interaction between client interface, server and third-party APIs (Mahmoud and Williams, 2016). This requirement is usually found relevant to the service providers and might be common in specific application domains like e-government and telecommunication. The implications of this requirement are usually not well understood during the system design (Ameller et al., 2016).

Reusability requirement is more relevant to the government and military application domain (Mairiza et al., 2010). Svensson et al. (2009) indicate reusability as one of the quality aspects that project leaders in their study reported important. Like maintainability, reusability is also described as one of the least important quality aspects during the software development (Ameller et al., 2016).

Maintainability

Mairiza et al., (2010) defines maintainability as *"capability of software product to be modified that many include correcting the defects or make an improvement or change to the system".* It is the ability of the system to minimize the ripple effect on the whole system when a small change is carried out. From Table 8, we can summarize that maintainability appear as a QR in five of the six articles considered. Despite the popularity in case studies, maintainability is categorised as one of the least important quality requirements which is usually neglected or dismissed during solution development (Svensson et al., 2012; Ameller et al., 2016). Adaptability, modifiability, compatibility are some quality attributes of the maintainability requirement (Poort et al., 2012).

Performance

Performance and usability are rated as two important quality requirements by companies (Svensson et al., 2009; Philips et al., 2012; Ameller et al., 2016). Performance requirements deal with the data flow or control flow specifications of the software solutions (Svensson et al., 2009). Performance is measured based on the response time of the system when an event is initiated. The events are messages, interrupts or requests from users to the system (Poort et al., 2012). The reason for the prioritization of performance requirement is highlighted in a case study as: *"performance requirements are always important since the software developed always acts as one sub-system in a larger system"* (Svensson et al., 2012).

The studies also suggest that performance requirements tend to impact several features of the software systems hence gaining its priority in companies (Svensson et al., 2012). The reasons that make the specification of performance requirements challenging are unclear rationality during requirement specification and the complexity in associating the performance requirements with data flow (Philips et al., 2012). 'Efficiency' and 'throughput' QRs are subsets of the performance requirement (Poort et al., 2012). The case study of Ameller et al., (2016) indicated the close relationship between 'dependability' and performance as QRs.

Privacy and Safety

Safety as QR has been initially classified for the safety-critical systems. Then safety as a quality requirement emerged mostly based on the application domain of the systems like energy resources, healthcare, transportation etc. (Mairiza et al., 2010; Mahmoud and Williams, 2016). Svensson et al. (2012) also pointed out safety as one of the prime quality requirements in B2B companies. Companies viewpoint on safety requirement is to make sure that no consumers, environment, or physical objects are harmed with the products they sell. The B2B companies specifically mentioned **'accuracy'** and safety as important QRs (Svensson et al., 2012). **'Privacy'** and safety as strong QRs of an experimental software system was documented based on the case study of a software system that controlled the users drinking habits (Mahmoud and Williams, 2016).

Reliability

Though usability and performance are mentioned as key QRs, reliability also got its position in the QRs list that companies portray as vital during software development (Svensson et al., 2012; Philips et al., 2012; Mahmoud and Williams, 2016). In general, reliability is defined as the capacity of the software to operate without any failure in the specified conditions during a given time period (Mairiza et al., 2010). Philips et al.'s (2012) case study on five companies also revealed reliability as one of the top three QRs that companies have their focus on. Some of the quality attributes that are closely associated with reliability requirement are accuracy, availability, recoverability, consistency, integrity and fault tolerance (Poort et al., 2012; Mairiza et al., 2010). Mahmoud and Williams's (2016) case study also exhibited the co-occurrence between **'accuracy'** and reliability in their case companies.

Security

Security requirements focus on resisting any unauthorized access to the system, program, and data (Mairiza et al., 2010; Poort et al., 2012). Svensson et al.'s (2009) study describes that product mangers rank security as the third most important QR during software development. In companies, 'integrity' and security requirements are associated together and used interchangeably (Svensson et al., 2009). Philips et al.'s (2012) investigation on five companies presents the participants viewpoint on internal quality with security as second most important QR. In the same study (Philips et al., 2012) '**accuracy**' emerged as the most important QR outweighing the security requirements.

Usability

Usability is a quality requirement that relates to the user interface specifications identified for the software solution and is one of the popular requirements in the companies (Svensson et al., 2009). It denotes how easy it is to use the software solution and the user support to tackle potential challenges. The challenges ranges as follows, reducing the occurrences of errors, incorporating user needs to design system efficiently, achieving the user satisfaction and the like (Poort et al., 2012). As this QR relates to the acceptance of the software solution with the end users, it receives highest importance in companies (Ameller et al., 2016). There is also an implication that the usability and performance requirements are not identified during the early stage or project planning stage of development which could lead to challenges in later development stages (Svensson et al., 2009).

Table 8 also signifies usability as one of the highly mentioned quality requirements. A case study highlights the quote from their interviewees i.e. *"if the product is not usable, we will not sell any products"* as the reason for prioritizing usability as a key QR (Svensson et al., 2012). In few companies, 'utilization' is listed as a QR which is used alternatively to usability QR (Philips et al., 2012).

3.3.4 Quality requirements of AI solutions

Quality requirements and its relation to AI is relatively new area of study, so there are not many articles showing its relation. There are different frameworks to classify key quality attributes and the quality attributes are continuously evolving (Glinz, 2007). This section focuses on the quality requirements that plays vital role in the design of autonomous vehicles (Cysneiros et al., 2018) and considering transparency as quality requirement (Leite and Cappelli, 2009).

Accountability

The software organisations should consider to involving ethics values or ethics requirements during the elicitation, modeling and analysis of requirements. The organisations should identify the ethical values and make sure the system or software they develop are aligned with those values (Aydemir & Dalpiaz, 2018). Accountability as ethical value sprung from ethical issues in software engineering. The ethical issue relating to accountability is: *"Who should be held responsible for the harm caused by software?"* (Aydemir & Dalpiaz, 2018). The other ethical value proposed like privacy is a quality requirement already. Likewise, accountability can also be updated as a quality requirement to ensure the quality of the system (Aydemir & Dalpiaz, 2018).

Privacy

Privacy is one of the key quality requirements. In autonomous system, privacy 'can and should be controlled by the individual users' (Cysneiros et al., 2018). The user should control how and where their personal data is utilised, and customise the data they are willing to share. The three types of privacy concerns that emerges because of autonomous vehicles are (Cysneiros et al., 2018):

- Personal autonomy
- Personal information
- Surveillance

Reliability

The AI or autonomous systems should be reliable for the users to accept the transformation to a new technology or AI solution. Cysneiros et al. (2018) points out that it is important to ensure robustness and capability of the system to handle daily operations like traffic flow, pedestrian movement, night-time movement and. For example, even the manufacturers and suppliers reliability should be verified in case of autonomous systems (Cysneiros et al., 2018).

Safety

Safety plays a significant role in influencing the adoption and acceptance of technologies. The need of safety is paramount which should tested by trial and error before advancing to implementation. Various practices and factors should be analysed and taken into account and the technology also should comply to the legal acceptance and liability (Cysneiros et al., 2018).

Security

Security is one of the well-known quality requirements in software engineering (Glinz, 2007; Mairiza et al., 2009). From the perspective of autonomous systems, security as quality requirement outlines different consequences and risks caused by the autonomous systems (Cysneiros et al., 2018). So, security as a quality requirement plays a significant role for developing autonomous systems.

Transparency

Transparency from the autonomous vehicle's perspective is to make it easy for the people to interpret what the vehicle is doing. The people should know why the vehicle makes certain decision which make them confident when using the autonomous system (Cysneiros et al., 2018). This helps in building the trust with the users and people. Leite & Cappelli (2009) and Cysneiros et al. (2018) advocate transparency as a key quality requirement for AI or autonomous systems. The software should show what information it used, why it is used and how it is used to form decisions (Cysneiros & Werneck, 2009).

Trust

The concept of trust exhibits its close connection with transparency. When the autonomous vehicles operate openly and transparently, it leads in gaining human trust (Cysneiros et al., 2018). Cysneiros &Werneck (2009) examine various perspectives like open software, software purity, cryptography, license and data collection with the goal to achieve and transparency and its connection with trust. Their results demonstrate the interdependency between transparency and trust, that is, in some issues trust either improve transparency or deter transparency and vice versa. In addition, trust as quality requirements also exhibits relation with other quality requirements like privacy, safety, reliability, and security (Cysneiros et al., 2018).

3.4 Summary

Ethical issues vs quality requirements

The different ethical issues that arisen because of the use AI or autonomous systems are transparency, security, safety, privacy etc. Based on the existing literature, the quality requirements that are popular in companies and quality requirements in AI are elaborated. By comparing the ethical issues and quality requirements some common factors between them are identified. Figure 6 illustrates the comparison on ethical issues of AI and quality requirements of AI.

The common factors are transparency, security, safety, privacy and trust. From the literature it is evident that security, safety and privacy are already treated as key quality requirements of software systems. The new quality requirements are transparency and trust which are found in both ethical issues area and in potential quality requirements proposed for AI. As Glinz (2007) mentions quality requirements of software system is evolving, transparency and trust manifest its potential to be key quality requirements.

The quality requirements accountability and reliability can also be the potential key quality requirements when we consider AI solutions. The quality requirements could differ between projects, but it is better to let the requirements engineers to decide and prioritise on that (Cysneiros et al., 2018).

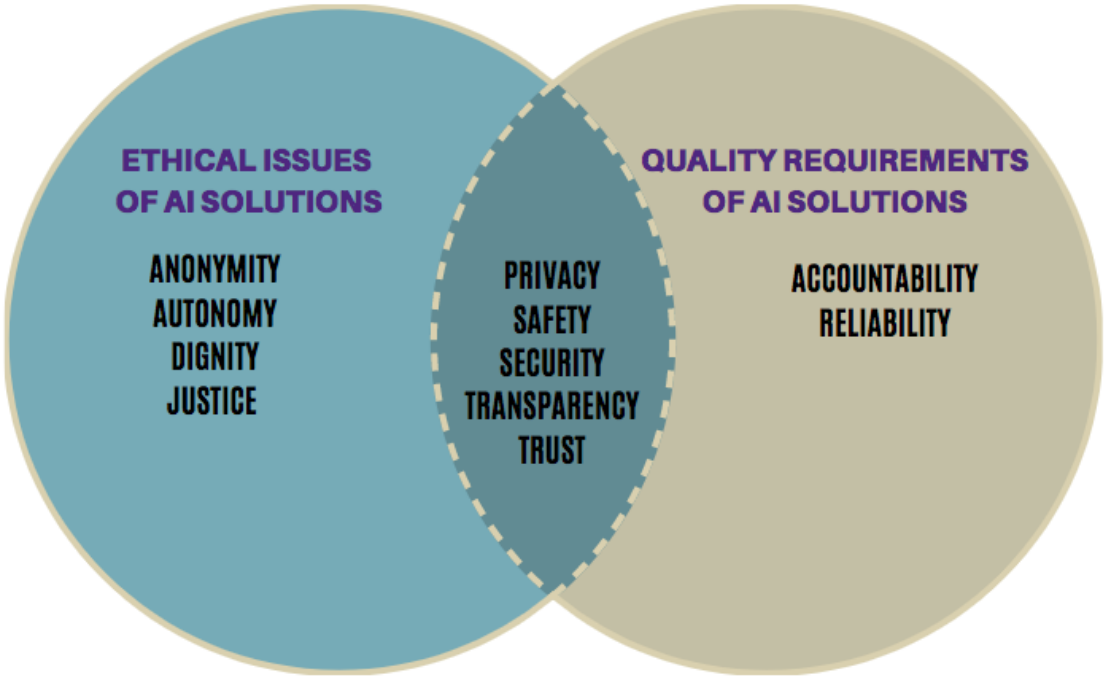

Figure 6: *Comparison between ethical issues and quality requirements of AI solutions*

4 Empirical Study

4.1 Ethical guidelines of AI solutions

4.1.1 Company A

Background

The key motto of the company is to adapt to ethical of way of working to provide ethical services and ethical banking. This company is one of the forerunners in Finland in applying AI and data science ethically. This company has also designed online courses which provide ethical principles and ethical software design coaching to their employees. Table 9 provides an overview of the ethical guidelines of Company A.

Accountability/Ownership

One of the ethical guidelines of Company A highlights the importance of accountability and ownership. The algorithms are made accountable by assigning owners to them and ensuring their ethicality. In the project, algorithms are assigned with dedicated owners who ensure ethical AI throughout the lifecycle of AI solutions. It is important that even when having the access to individual level data, the identity of the customers should not be traced which also helps in keeping the privacy of the customers intact.

Correctness

Company A stresses the correctness of data used and the correctness of results produced by AI. When wrong data about the people are used then the decision made are likewise wrong decisions. The wrong decisions proposed for customers leads to influence in the trust factor. “Is the data concerning the customers, right?” is one of the fundamental questions that has to be analysed and the correctness of the data must be ensured when developing AI solutions. So, when developing AI, the correctness of data has to be ensured which also influence the correctness of results.

Fairness

Fairness has not been explicitly defined as an ethical guideline for use of AI but holds an important place in the company’s way of working. Fairness is an ethical way of working to assure trust between the customers and the organisation. An important question from the customer perspective that should be dealt when handling AI projects is *‘Can I trust that the analysis and decision concerned me are made correctly by following the correct process?’.* Acting fairly is the key point when handling this certain question. When the decisions made are fair, customers stay loyal and trust the services delivered by the company.

Privacy

When considering privacy, it is paramount to safeguard the personal data in the vast available data about the customers. This eliminates pinpointing any person when analysing the data. It is also important for the company to adhere to data protection policies and to provide an extensive data protection education to the concerned employees. The company has online courses on privacy, privacy issues, GDPR utilisation, data security, privacy regulations and special financial services education. This training ensures making better analytics with the customer data.

With the GDPR and privacy regulations in action, the first factor that should be recognized before making any analytics is: *'What is our right in utilising the data?'*. It is essential to think from the customers' perspective. According to the laws of privacy, customers are foremost and they possess complete right to know and control what is done with their data. From the available data it is crucial to differentiate the common data and conflicting data for creating analytics. With an access to the conflicting data, employees have the potential to *misuse* data which leads to the risk of customer privacy.

Transparency and Openness

Company A highlights transparency and openness together as one of the ethical guidelines devised for the use of AI. To attain the goal of transparency, operating openly with the customer and stakeholders is seen vital. It is also important for the company to be open about the way they utilize AI in their projects. The guideline portrays an interrelation between openness and transparency when designing AI applications.

From the company perspective various factors like privacy, ecology, sustainability, social responsibility, social equality, and human rights. are seen ethical. The interviewee highlighted that, it is necessary to work ethical by being open and equal with the customers, stakeholders and partners. The checkpoint needed for being open and transparent is to make sure that analytics and data are right.

Trust

From the company perspective, trust is associated with various AI ethical guidelines and key factors like accountability and correctness. For all the financial and banking service providers, doing ethical analytics is important because they provide essential services to the society. It is a basic principle for banks to reach the level of trust and fulfil the users. Another viewpoint emphasized is 'internal trust' which ensures ethical way of working inside the company. The employees must be aware of what they are supposed to do with sensitive data and follow the ethical way of working promoted in the organisation.

Trust which is a core concept have to looked from the customers' perspective by taking into account of the following questions:

- *'Can I trust that data concerning me is right?'*
- *'Can I trust that the bank is operating in fair manner?'*
- *'Can I trust that the information that concerned me stays private?'*
- *'Can I trust that the analytics and decisions concerned me are made correctly?'*

These questions are used to look into various ethics related concerns that might arises with AI utilisation. From company's perspective, these questions relate to privacy, fairness, correctness, openness which influences the trust factor directly. Any compromises made during analytics leads to people considering AI as threat. AI and data science have to work ethically to gain people trust and put that in right direction for AI development.

Table 9: Overview of the ethical guideline in Company A

Ethical guideline	***Outline of the ethical guideline***	***Traceability of the guideline***
Accountability/ Ownership	Assign owners to all the principles and algorithms that we develop	Interview & document
Correctness	Check whether the data concerning the users are right	Interview
Fairness	Act fairly and follow the right process when making decisions	Interview
Privacy	Safeguard the protection of privacy and personal data in accordance with data protection policies	Interview & document
Transparency and Openness	Operate openly in relation to our customers, partners and stakeholders and ensure transparency of AI application	Interview & document
Trust	-Ensure there is internal trust when working -Make sure there is no compromise made in the analytics of user data	Interview

4.1.2 Company B

Background

Case company B is one of the biggest trading and retail companies in Finland. Sustainability is one of the driving factors of the company. As a part of their sustainability goals, adopting ethical use of AI had been included recently. AI is used in the company to help them tailor benefits for the customers by matching customers preferences and interests. It is envisioned that use of AI will help them provide best individual services for millions of customers. The probing question that is helpful when utilising AI is: *'Are we doing the decisions as our guidelines and principles that we have promised?'*. Table 10 provides an overview of the ethical guidelines of Company B.

Data

When creating AI services that makes personalised recommendations to the customers, it is important to understand the data in the first hand. Understanding the data to be used by AI helps in ensuring transparency and openness of the AI. For example, the current AI service provides personalised offers and benefits of products through channels like mobile apps and e-mails. To make ethical choice of products to be displayed in the AI solution, the answer for the question *'What kind of data we will use for this (choosing products)?'* need to be identified.

When the partners develop the AI systems of the company, the discussion about data have to be made with them to understand the black boxes. And discuss further about being responsible on how the data is utilised in the system. From the perspective of the company, the questions that have to be discussed with partner organisations and external vendors who provide AI services are:

- *'What kind of data are they using?'*
- *'How are they developing the system?'*
- *'How to proceed further if some information is found missing?'*

Fairness

The use of AI should yield common good for the society and improve the welfare of the Finnish society. The guideline for acting fairly has been described as: respecting the human rights and avoiding discrimination. The use of AI should not lead to discrimination of users of the service. This relates to one of the questions of the company's strategic action, that is: *'how we work with the data?'* which ensures responsible AI solutions.

Purpose driven

When developing AI services, customer data is used only for the purposes that customers have granted permission for data utilisation. Customers can decide whether to provide complete data access or regulate what data that the company can use. When AI is purpose-driven, the checkpoints of how

to be trustworthy and how to adopt customer perspectives are taken into account. This helps in having a holistic approach where not just the technical perspective or technical issue is focussed.

The ethical principles help in addressing the technical perspective in designing a trustworthy AI. According to the ethical guidelines of the company it is also important to consider the non-technical perspective by being purpose-driven and include customers perspective when designing the AI services.

Responsibility and Security

Responsibility and security have been focused in the first ethical principle for using AI. As per the principle it is important to protect the *privacy* and data of the customers. Being responsible on AI usage and ensuring security are the initial criteria that help in data utilisation and creation of AI solutions and algorithms. As a part of developing responsible AI, it is important for the team to understand data utilised when designing AI services. In addition, to employ responsible AI two key questions that should be transparently answered to the customers are:

- *'Where is the user data utilised?'*
- *'How is the provided user data being utilised?'*

Answering the above questions is the initial stage where the data scientists can discuss about how to explain the system to the users.

Sustainability

Sustainability is portrayed as one of the key values of Company B. Sustainability strategy is devised every year and AI ethics has been a part of the strategy since spring 2017. The company has a smart shopping application that supports customers making ethical choices by listing products for achieving healthy lifestyle. The viewpoint of 'AI as Service' has been proposed for helping the customers to make right choices. This helps in providing customer-specific choices and recommendations of products. To provide personalised recommendations for customers and help them choose right things, two questions need to be answered:

- *'How to support customers to make sustainable choices?'*
- *'How to support customers to live healthier?'*

Based on the customer needs and goals, the question of *'Which product to promote?'* out of million different products available is where the potential of AI will come into play. That is, using AI to understand the available products and customers viewpoints and interests.

Transparency, Openness and explainability

Being transparent in utilising the user data is already indicated with the responsibility principle of AI utilisation. The current objective of the company is to probe how to inform that AI is behind the recommendations shown in their AI service. It is also highlighted that it is vital to be transparent by showing AI behind the service they are using. The interviewee highlighted that it is important for the data scientists to have skills to explain the operations and reasons behind the AI. It is like opening the black box and discussing what is inside it and this helps in understanding the analytics better. By

having this kind of discussions with partners who offers AI systems is important for transparent working and better understanding of AI.

Handling massive amount of customer data leads to various ethical questions during AI service development. To handle this, the first basic question that have to be considered is: *'What data are we using?'.* This fundamental question is strengthened by the following subordinate questions like:

- *'What data is available and what kind of data are we allowed to use?'*
- *'What is the privacy?'*
- *'What kind of rights do we have on data?'*
- *'How we get the license?'*

Answers to these questions help in explaining the customers transparently about their data and identify how to improve service and create new services from it.

Trust

Designing a trustworthy AI is represented in association with the ethical principles of AI solutions. By being transparent with the decisions made and designing AI based on the purpose of AI are the cornerstones for achieving a trustworthy AI. When the company acts open and transparent with the customers, the trust on the company and their products increases. When the customers trust the company, they share their future data with them.

Table 10: Overview of the ethical guideline in Company B

Ethical guideline	***Outline of the ethical guideline***	***Traceability of the guideline***
Data	-Discuss how the data is utilised in the system -Understand what user data is available	Interview
Ethical design	Ensure better decision making and best customer experience	Interview
Fairness	AI should avoid discrimination of users	Interview
Purpose-driven	Create solutions that are useful for our customers	Interview & document
Responsibility and Security	-Protect data and inform users transparently on how and where the data have been utilised -Responsibility and security should direct the collection and utilisation of data	Interview & document
Sustainability	Support customers to make right and sustainable choices	Interview
Transparency, Openness and Explainability	-The decisions made by AI should be understandable and explainable to users -Explain customers on what data is being used and be open	Interview
Trust	Operate openly and transparently with the customers	Interview

4.1.3 Company C

Background

The company values which form as their heart of working are: *Trust, Caring, Transparency and continuous leaning.* The company provides support to their clients to tackle the ethical implications and concerns relating to the AI systems. There is a mandatory 'ethics in data science' or 'AI ethics' training given to their employees. The ethical principles were formulated as a tool for decision making when building autonomous systems. The ethical principles have been designed based on the company values. The company has also designed data ethics canvas with a set of questions to identify potential ethical issues and the actions needed to mitigate the issues.

Accountability

Transparency and accountability are documented together in the company's data ethics canvas. According the data ethics canvas, each project has a person that takes responsibility in situations where the system is causing harm. This helps in mitigating the risks as soon as it appears instead of finding an accountable person in the later stages. The set of questions relating to accountability in data ethics canvas are:

- *'Who will be responsible and contacted if this system/product/services seen causing harm?'*
- *'Who in the project resolves the ethical issue?'*

Inclusion and Fairness

Inclusion in data is important to understand who are the diversified set of people that will be impacted by the AI solution. All those people who will be impacted should be represented in the data sets. The unfair outcomes are possible with biased data sets, so it is crucial to avoid creating and reinforcing bias. Though it is highly context sensitive, there is no best practices to handle biases. There are infinite number of ways to tackle fairness related ethical issues, for example, considering different perspectives to make the system less biased. By using diverse set of test data both fairness and inclusivity is addressed.

From the company perspective, inclusion in the team is seen vital. To resolve issues by considering various angles and reasons, the projects need to have multi-disciplinary teams. It is important to hear different viewpoints that are relevant to the project. As a part of quality work, inclusion of multi-disciplinary development teams should be highlighted to the clients by explaining the need of it.

Privacy and Safety

The AI systems can use data which also includes personal data of the users. This introduces the vulnerabilities which demands to be extra careful with privacy and security. By default, privacy and safety of the personal data have to be ensured. As a step to ensure privacy and safety, it is practical to inform the users what kind of personal data are used in the system. It is also better to collect as

fewer sensitive data about the users. This can be held by anonymising as much data as possible. In the ethics canvas, the questions relating to safety are:

- *'Can the system be 'played''?*
- *'How are ensuring the data safety?'*
- *'Is it possible for the malicious actors to intentionally input bad data in the system?'*

Purpose and Impact

When designing an AI system, the purpose of the system and the people who will be affected should be considered. The positive and negative consequences of the system have to be identified for the user groups and the extended target users. The related questions for the measuring the impact in data ethics canvas are:

- *'Possible positive effects?'*
    - *'For whom?'*
    - *'How to increase them?'*
    - *'How to monitor/measure?'*
- *'Possible negative effects?'*
    - *'For whom?'*
    - *'How to minimise them?'*
    - *'How to monitor/measure?'*

The data ethics canvas of the company highlights the analysis of the purpose and impact of AI solutions. In the data ethics canvas, the concept of purpose and impact are categorized in a category called *intent*. It takes in a detailed viewpoint by considering various factors that will be affected by the AI system that is been built. So, the set of questions that have to be considered relating to the purpose or intent of the system are:

- *'Who does the system you build affect?'*
    - *'Intended audiences?'*
    - *'Externalities?'*
    - *'How much uncertainties can be tolerated in terms of different stakeholders?'*
    - *'Who will be affected and how by false positives/negatives?'*
- *'What are the expected results of the system you are designing/building?'*
- *'What is the state you will be measuring against?*

Transparency and Explainability

The company strives to adhere to transparency in all their projects. The transparency criteria are taken into account from the very first stage of the project and it directly impacts on the model they choose for development. Neglecting the transparency principle in the beginning stage would lead to unpleasant experiences in the later stage of development. The ethical principles do not say what should be done exactly but it acts as guide which supports in auditing and dealing with ethical issues.

When discussing transparency from design perspective, it should be understandable rather than just focusing on convenience of use. The balance between transparency aspect and convenience aspect have to be balanced out by designing and iterating. The true challenge of understanding AI arises when it is hard for the people who designed the AI system to explain the system. For example, when dealing with a sensitive project say public service system, the expectations on explaining the system will be high. In such cases it is crucial to select a model that is accurate rather efficient. So, there occurs some essential trade-offs which need to be addressed in the beginning stage of development.

When considering transparency in a project, there are situations when there is no direct access to the information used in analytics. In those cases, it is impossible to explain the AI system without knowing the kind of information. So, transparency is more about *access* and *intelligibility*. Moreover, transparency in AI systems does not only mean the technical transparency of exposing data and algorithm, but also specifies the explainability and understandability of the system. The questions relating to transparency and explainability which helps in identifying ethical concerns are (from data ethics canvas):

- *'Is explainability needed?'*
    - *'Who needs it?'*
    - *'How much are you able to explain your system?'*
- *'Are there any legal issues, regulations or policies that should be taken into account when using data?'*
- *'How to make sure that users exercise their data rights?'*

Trust

The ethical principles designed for using AI shows trust on AI systems relating to the transparency concept of AI. A trustworthy AI is ensured when the purpose of the system is clear and behaves as expected. When the AI system is adhering to the ethical principles it paves the way for achieving trust and thus satisfying one of the company's value i.e. trust. Avoiding distrust is also encountered in the general guidelines for ethical AI which advocates not to promote division in the society.

Table 11: Overview of the ethical guideline in Company C

Ethical guideline	***Outline of the ethical guideline***	***Traceability of the guideline***
Accountability	Mitigate risks by assigning accountable person to handle and resolve risks	Interview
Inclusion and Fairness	Aim for inclusion by striving to understand who are all the people that will be impacted by the system that we design and build	Interview & document
Privacy and Safety	Collect, store and use personal data safely and default to high privacy	Interview & document
Purpose and Impact	Focus on purpose and impact over technology when developing AI	Interview & document
Transparency and Explainability	Prioritize transparency in the system that we design and build and strive to increase trust in all of AI solutions	Interview & document
Trust	Adhere to ethical guidelines of AI and emphasize transparency	Interview & document

4.2 Ethical practices when developing AI solutions

The analysis of interview data revealed design practices to be used in the case companies during the AI system development. They were highlighted in the interviews as way of recognising diversified viewpoints when building the system.

4.2.1 Using multi-disciplinary development teams

The interviewees pointed out in having multi-disciplinary team which included experts like data scientists, designers, and sociologists. It is a practice to be incorporated during solution design. The interviewees mentioned the following two practices regarding the usage of multi-disciplinary teams when developing AI solutions:

- Include multi-disciplinary teams in the project to ensure quality work.
- Establish teams with different skills, capabilities and background.

This approach ensures the quality of the AI solution and helps in assessing how the AI solutions will impact people. Moreover, multi-disciplinary teams bring in value by identifying and resolving ethical issues from different perspectives when designing AI solutions. An interviewee's view was:
"the key thing I found is that, how to make sure that we are truly in turn multi-disciplinary. I think it's the key of it all. We need different competencies; we can't only have data scientists."

4.2.2 Defining the purpose and impact of AI solutions

The companies are aware of impact made by the new technologies and have emphasized that as one of their ethical guidelines. The set of guidelines that the companies proposed in order to track the impact of the applications are as follows:

- Identify all the people that are influenced by the application or technology.
- Formulate a clear purpose or resulting behavior of the AI solution.
- Define potential positive and negative effects or consequences of the solution in early stage.
- The process of application design has to be purpose-driven.

It is essential for companies to focus on the people or target users who will be impacted with the designed application. In addition, for AI applications, measuring the impacts beyond target groups is significant. *"so, focusing on the purpose and impact, … think about people first, think about all the people that might be impacted, also beyond your own target groups".* Table 10 and Table 11 which gives the overview of ethical guidelines in company B and company C respectively shows purpose (and impact) ethical guidelines and their traceability.

4.2.3 Communication with customers and people-oriented approach

When developing AI system getting users or customers feedback is important in knowing the customers perspective which can be utilised to design ethical system. One of the interviewees mentioned that getting customer feedback assists in transparent working but the ways to communicate with users is still unclear. An interviewee mentions it as: *"To be honest it is quite difficult to think that, what is right and how we should communicate it to users, because it is difficult task"*.

By utilizing AI, it is important to place the interest on the customers. As the AI is used to facilitate people's everyday life, the goals have to be precise and user needs should be the priority. When performing data analytics, customers perspectives is taken into account first. Customers should have power on their data and control what information that the company can utilize.

4.2.4 Ethical design

The main objective of designing an AI service is to provide best customer experience. The service should help in better decision making and improving efficiency. In cases like providing personalised recommendations, the service needs good testing and validation before going to the production stage. The algorithms used for suggesting products should also be improved based on the feedback from the users. As mentioned earlier in the sustainability section, AI can be a service assists in making choices

based on the user goals. Table 10 which gives the overview of ethical guidelines in case company B outlines the 'ethical design' guideline devised in the company and its traceability.

4.3 Summary

RQ2: Ethical guidelines for using AI

This section gives an overview of second research question: ***What kind of ethical guidelines do the case companies have for designing AI solutions***. After analysing the different ethical guidelines mentioned in the documents and interview, the different ethical guidelines in the case companies are documented in the previous sections. The companies in this case study have designed guidelines for ethical issues which they consider relevant for their domain. The ethical guidelines between the companies had some commonalities. Table 12 presents an overview of the ethical guidelines of AI designed in the case companies.

The ethical guidelines on fairness, transparency and trust are defined by all three case companies. Furthermore, the companies had more than one ethical guideline planned to tackle the ethical issues which are more relevant to their domain. For example, the guidelines like 'working together with users, customers and stakeholders, and being open and the ability to explain the solution to the users' are developed by case company B to address the transparency related ethical issues.

In addition, there are other ethical guidelines proposed by the interviewees which are not explicitly related to any ethical issues but are suggested as ethical practice when using AI. Hence, they aren't categorized in Table 12. Those guidelines are aimed to enhance the quality of the designed AI solution. For example, the guideline 'Use multi-disciplinary development teams when designing AI solutions' as deduced by companies aid in identifying potential ethical issues from diversified perspective and resolve ethical concerns during solution design. And the guideline on reporting the purpose of AI solutions helps in comprehending the purpose of designing a solution. This also helps in knowing the potential positive and negative impact that AI solutions could bring upon to the society and the environment.

Table 12: Overview of ethical guidelines of AI in the cases companies

Ethical guideline	*Outline of the guideline*
Accountability (Company A and Company C)	- Assign owners to all the principles and algorithms that we develop. - Mitigate risks by assigning accountable person to handle and resolve risks.
Correctness (Company A)	Check whether the data concerning the users are right.
Fairness (Inclusion) (Company A and Company B: Fairness; Company C: Fairness and Inclusion)	- Act fairly and follow the right process when making decisions. - AI should avoid discrimination of users. - Aim for inclusion by striving to understand who are all the people that will be impacted by the system.
Privacy and Safety (Company A: Privacy; Company C: Privacy and Safety)	- Safeguard the protection of privacy and personal data in accordance with data protection policies. - Collect, store and use personal data safely and default to high privacy.
Security and Responsibility (Company B)	- Protect data and inform users transparently on how and where the data have been utilised. - Responsibility and security should direct the collection and utilisation of data.
Sustainability (Company B)	Support customers to make right and sustainable choices.
Transparency, openness and explainability (Company A, Company B and Company C)	**-** Operate openly in relation to our customers, partners and stakeholders and ensure transparency of AI application. **-** The decisions made by AI should be understandable and explainable to users. **-** Prioritize transparency in the system that we design and build and strive to increase trust in all of them.
Trust (Company A, Company B and Company C)	- Ensure there is internal trust when working. - Make sure there is no compromise made in the analytics of user data. - Operate openly and transparently with the customers. - Adhere to ethical guidelines of AI and emphasize transparency.

RQ3: Ethical guidelines and quality requirements

The ethical guidelines defined in the companies for using AI and the quality requirements identified from the literature signifies potential relationship. Table 13 shows how the ethical guidelines discovered from the case companies relates to the quality requirements. The three sub categories of literature study findings on quality requirements are: 1) Mairiza et al.'s (2010) article which summarizes over 100 quality requirements found in scientific publications; 2) quality requirements highlighted in the case study articles published in last 10 years and; 3) quality requirements of AI solutions which are mainly based on an ongoing research about quality requirements in autonomous systems.

Table 13: Relating ethical guidelines of AI with the quality requirements

Empirical Study Results	**From the literature study**		
Ethical guideline for using AI	**Quality requirements (Mairiza et al. (2010))**	**Quality requirements highlighted in case studies (Section 3.3.3)**	**Quality requirements of AI (Section 3.3.4)**
Privacy	X	X	X
Safety	X	X	X
Security	X	X	X
Accountability	X		X
Trust	X		X
Correctness	X		
Sustainability	X		
Transparency			X
Explainability			
Fairness			

Table 13 is formulated to compare the ethical guidelines of AI and quality requirements. This objective of relating the ethical guidelines of AI with quality requirements comes from the literature which aimed to define quality requirements of AI solutions (Cysneiros et al., 2018). Privacy, safety and security ethical guidelines are appearing in-line with the quality requirements. This exhibits the

connection between ethical guidelines of AI and quality requirements. The ethical guidelines can be utilised for identifying and defining quality requirements of AI solutions.

Accountability, trust, correctness, sustainability and transparency ethical guidelines also appears in the literature as quality requirements. There are some new candidates like fairness and explainability which are not mentioned as quality requirements in the existing literature. But fairness and explainability ethical guidelines are defined by the case companies highlight their importance when building AI solution. So, the ethical guidelines on explainability and fairness can be utilised to define the explainability and fairness quality requirements during AI solution design.

5 Discussion

This section discusses the answers to the research question and compare them with existing research knowledge. We also discuss the limitations of the study.

5.1 RQ1: Ethical issues and quality requirements of AI solutions– Relationship

The first research question that is answered during this study is *"What is the relationship between ethical issues and quality requirements of AI solutions?".* The analysis of evolution of ethics is used as a basics to highlight the trend in ethics. The literature study on ethical issues of AI solutions was taken as basis to spotlight different ethical issues discussed in the literature and the importance of considering it when designing AI solutions. The goal was to show the ethical issues mentioned in existing literature, and explore the relationship between ethical issues and quality requirement of AI solutions.

Ethical issues of AI

The research studies delineate the different ethical issues that had been arisen or might arise of AI in future. The key categories of ethical issues identified from the existing literature are: privacy, autonomy, anonymity, transparency, security, safety, justice and dignity. These ethical issues categories exhibited relationship with trust which is seen as the core concept relating to the ethical issues. The trust on the AI solution is directly affected by the ethical issues and that signifies its relation with different ethical issues.

Quality requirements

Quality requirements denote to the concerns and attributes that are not related to the functionality or feature of the software system (Chung & Leite, 2009). Quality requirements are defined based on the software system and it is essential to define them during system design. QRs impact when making the design decisions and architecture alternatives of the software system (Mahmoud and Williams, 2016). As the quality requirements help in verifying the quality of the software system build it is crucial to consider QRs during system design and development.

The quality requirements if not handled properly might lead to failures or most expensive changes in software development (Svensson et al., 2008). The analysis of articles about quality requirements employed in the case studies showcased some quality requirements that are seen significant or prime. The quality requirements that are highlighted in the case studies are: usability, performance, reliability, security, safety, maintainability, accuracy, interoperability and reusability. This list represents quality requirements that companies consider critical when developing software systems or solutions.

Cysneiros et al.'s (2018) study highlights transparency, privacy, security, safety, reliability and trust as core requirements when building autonomous systems or deploying AI in systems. The role of

these core requirements as quality requirements are argued in this study. The interdependency between the quality requirements like transparency and trust; safety and security; impact of transparency on privacy are critiqued to signify how the quality requirements might positively or negatively influence one another. QRs like security, privacy, safety and reliability are already in the focus when developing software in companies. As the impact of AI solutions are both in organisational and societal level the importance of defining the quality requirements of AI solutions takes its spotlight. In my opinion, these quality requirements are crucial when developing autonomous systems because satisfying these quality requirements builds confidence and trust, and enhances the customer experience.

Relationship between ethical issues and quality requirements of AI solutions

The reason why ethical issues of AI are identified is to show the different ethical issues that are identified impactful in the societal level. The ethical issues of AI solutions and the quality requirements identified as important when utilising AI exhibits an explicit interrelation between them. The categories of ethical issues like transparency, security, safety, privacy and trust are highlighted as vital quality requirements for building AI or autonomous systems (Cysneiros et al, 2018). This indicates the role of quality requirements during the AI system design, in order to identify and eliminate the possibility of ethical issues in the developed system.

The other ethical issues categories, autonomy and anonymity are subordinately related to privacy. So, elaborately discussing the privacy requirement during AI solutions development would establish a possible early check on identifying ethical issues. Furthermore, justice and dignity relate to human factors of treating without any discrimination based on data and respecting human dignity which are not explicitly relating to any of the quality requirements. But the issues on justice i.e. handling data without discrimination relates to acting fairly. Fairness when designing could be a quality requirement in the future which would be quintessential for AI solutions.

5.2 RQ2: Ethical guidelines for using AI in case companies

The second research question that was answered during this thesis is *'What kind of ethical guidelines do the case companies have for designing AI solutions?'*. The goal of this research question is to explore the AI ethical guidelines used by companies to tackle any prospective ethical issues in the AI solutions. The results of studying ethical guidelines of AI helped in uncovering the perspectives of each case companies on their ethical guidelines of AI solutions. There were some recurring themes and also there were company-specific guidelines which are defined based on the company values or key ethical issues.

- **Accountability/Ownership:** The AI ethical guideline on accountability is all about having accountable persons or owners to deal with ethical issues and ethical concerns that arise when designing AI solutions. The person who is assigned accountable should take responsibility for resolving the issues and make sure that the developed AI solution is ethical throughout its lifecycle. This accountability guideline helps in reducing the confusion of the project members on whom to contact when an ethical issue or concern emerges during the design.

- **Correctness:** The user data are utilised to form results, recommendations or decisions in the AI solutions. So, correctness as ethical guideline ensures correctness of the data used to make the decisions. The companies should also verify the correctness of the result to ensure there is no wrong recommendations or results given to the use

- **Fairness:** The ethical guidelines of fairness are represented on both company perspective and AI solution perspective. Fairness in the company perspective describes about ethical way of working to ensure customer trust and loyalty. In addition, it also helps in avoiding discrimination which would impact the society in a positive way. From AI solution perspective, to ensure fairness, it is vital to check up on the inclusion of data. In order to reduce the discrimination or bias on data used to make analytics, it is necessary to include diversified data sets representing every possible target groups.

- **Privacy, Security and Safety:** The reason for grouping these three guidelines is because they are primary guidelines not only for the AI projects but also for any software projects developed in companies. The personal data safety is a key criterion for ensuring safety, security and privacy of the customers. It is also crucial to abide to privacy rules and data protection policies to make sure about the level of rights that companies have on available user data. The key techniques that can be employed to ensure privacy and safety are: 1) Avoid collecting sensitive data; 2) Anonymise user data to avoid traceability of users.

- **Purpose and Impact:** The ethical guideline of purpose and impact ensures whether AI solutions developed adhere to the purpose of their creation. By checking the purpose of AI during the AI solution development avoids any potential deviation that could come up when playing with the customer data. It has to be ensured whether data is utilised only for the purpose the users have given permission and the company's right on data.
  When documenting the purpose of the AI system, the impact it creates on their user groups should be identified. This helps in identifying the potential positive and negative impacts that can occur with the system during the initial stages of development.

- **Responsibility:** The case companies talks about developing responsible AI. But the objective behind creating responsible AI is vaguely defined by the case companies. The guidelines in the company for developing responsible AI relates to various other guidelines like privacy, security and understandability. When developing responsible AI, the developers feel responsible for tackling the ethical concerns emerged.

- **Sustainability and Ethical design:** The AI solutions designed in the companies should assist the users to make right choices. These right choices may also have influences in the societal level. The design of AI services should provide best customer experience and also make right choices. So, the AI service should be verified and validated and utilise the user feedback to improve the algorithms.

- **Transparency, openness and explainability:** The ethical guidelines of transparency relates to openness and explainability aspects of the AI solution. By openness as guideline it means that the user data applied to form results and suggestions should be openly available. This helps the users to see what data of them is used by the AI solution. This guideline enables the companies to meticulously check whether the user data utilised is right and whether the analytics is done right. Moreover, the transparency guideline indicates the openness of data and algorithms.
  The key aspect of transparency guideline is to make sure that customers understand that AI is behind the results. In addition, the data scientists must be able to explain how and why certain decisions are made. When the transparency guideline is compromised in the project it leads to unpleasant experience.

- **Trust:** Trust is commonly associated with other AI ethical guidelines like transparency, accountability, correctness, openness and fairness. Trust is seen as a key factor for them to satisfy the users and maintain cordial relationship with their users. As data is key for most AI services, the users for giving their future should really trust the companies. Thus, when the AI solution adheres to the AI ethical guidelines committedly it leads in ensuring the trust factor.

5.3 RQ3: Using ethical guidelines to define quality requirements of AI solutions

The results of this thesis shed light on the different ethical guidelines that are devised in companies when designing and developing AI solutions. The ethical guidelines are: accountability, correctness, explainability, fairness, privacy, safety, security, sustainability, transparency and trust. The comparison with the existing literature indicates that the ethical guidelines highlighted by the case companies exhibit some connections with ethical issues and quality requirements identified in literature.

Ethical guidelines and quality requirements

The results also revealed that the case companies pointed out ***accountability*** as a guideline to tackle and deal with ethical issues. Assigning a responsible person will assist in monitoring and identifying potential ethical issues in every stage of the system development. Jones et al. (2015) and Royakkers et al. (2018) pointed out that responsibility and accountability should be in focus when incorporating artificial agent in the systems to track how the data is utilized. Furthermore, in software engineering, accountability is categorized as a key ethical issue that impacts quality (Aydemir and Dalpiaz, 2018). Information accountability intertwined with transparency is stated as a quality requirement of software systems (Leite and Cappelli, 2010). Likewise, considering accountability as one of the quality requirements when designing AI solutions requires more study and this will be investigated in the future research. The ethical guideline relating to accountability can be utilised in defining accountability as quality requirement of AI solutions.

Correctness as ethical guideline focuses on checking whether the data relating to the users are correct. This ensures the results generated based on the user data is also correct. Mairiza et al. (2010) article shows correctness as one of the quality requirements when designing software system. The AI solution should utilise right data in order to carry out analytics and provide user-specific result or recommendations. So, the ethical guideline relating to correctness can be used to for eliciting quality requirement that will help tackling the correctness related ethical issues when building AI solutions.

The interviewees of the case study referred to the concept of ***fairness*** of data. This is to avoid the inclusion of biased data which lead to discrimination. Training AI with unbiased data sets helps in making fair decisions and good quality solutions. Similarly, Royakkers et al. (2018), and Bostrom and Yudkowsky (2011) reported the ethical issues that occurred because of wrong judgements of AI solutions. Any intended or unintended exclusion of data could produce unfair results. In some cases, values like justice and integrity are violated by categorizing data to provide convenience for users of certain applications (Royakkers et al., 2018; Bostrom and Yudkowsky, 2011). An example is profiling jobseekers based on their gender and race (Royakkers et al., 2018). This influences the users' trust in those applications or services. Fairness can be the quality requirement specific to AI solutions. The ethical guideline of fairness in the companies can be used when defining the fairness quality requirement.

The next ethical guideline that is focused is ***privacy***. The viewpoints of the interviewees are to ensure the protection of user data and to eliminate any potential manipulations by design. The lead designer we interviewed pointed out the difficulty in design to change privacy in current services. Our findings also revealed the following perspectives on privacy implementation at companies. Firstly, it is important to discuss the scope of privacy and user data that have to be utilized by AI solutions. Secondly, making sure the level of user data that a company have the right to exploit. Thirdly, checking on what kind of user data is collected to avoid protecting volumes of sensitive data. Royakkers et al. (2018) pointed out that a strict surveillance of privacy when designing AI solutions is needed to track how data is used, how data is protected and identify possible misuse of data. Moreover, privacy as a quality requirement emphasizes on user data protection and users' right to control their data (Aydemir and Dalpiaz, 2018; Cysneiros et al., 2018, Zinovatna and Cysneiros, 2015). The overview of ethical guidelines on privacy by companies converges with the stance of privacy as a quality requirement.

Privacy is eminently related to ethical issues like security, safety and trust. As the AI solutions utilize and handle large volume of data, the security mechanisms to protect those data have to be intact. The guidelines propound to ensure the users' safety by extrapolating the threats with AI solutions when mishandled. The guidelines are proposed to ensure ***security*** and ***safety*** of AI solutions. Security and safety are quality requirements when designing any system or service. Cysneiros et al. (2018) already pointed out security and safety as important quality requirements of autonomous systems. Similarly, the ethical issues and guidelines relating to security and safety in AI solutions outlines their link when defining the quality requirements.

Sustainability as ethical guidelines of AI focuses on helping the users to make right and sustainable choices. The right choices should also align with the customers or users' goals and needs. Mairiza et al.'s (2010) list portrays sustainability as a QR. In addition, the recent study by Duboc et al. (2019)

highlights the importance of sustainability when building software systems. Requirements engineers should document the consequences of the system by considering the sustainability factors. Though RE is amiss when developing software solutions, the importance of considering the impact on wide angle is critical, especially from the sustainability perspective (Duboc et al., 2019). This shows the potential of having sustainability as a quality requirement when developing AI solutions to identify its impact on the society, environment and economic factors.

To look more in detail, the results of the study showed ***transparency*** as one of the key ethical guidelines for utilising AI. All of the interviewees agreed on the demand of including transparency as an ethical criterion when designing AI solutions. They described that the black box nature of AI solutions masks the type and quality of data used to train AI solutions. The guideline of being open to customers and stakeholders on customers' data utilization also aims to address the privacy related issues. The results suggested by AI solutions should also be explainable i.e. data scientists should be able to explain how the results are formed.

Explainability is related with transparency ethical guideline in the case companies. Customers possess the foremost right on their data and they could regulate how their data is represented in a system. The studies of Leite et al. (2010), Hoesseini et al. (2018), and Cysneiros et al. (2018) report transparency as a key quality requirement when developing information systems or automated decision-making systems. Leite et al. (2010) and Cyneiros et al. (2018) indicated transparency as one of the quality requirements of AI solutions. Transparency in software system is mentioned as one of the key factors to measure the quality (Hoesseini et al., 2018). The ethical guideline relating to transparency and explainability portrayed by the case companies can be used in defining the quality requirements of AI solutions.

Jones et al. (2015) have also stated that issues relating to transparency in decision-making systems arise because of the invisibility of data used. In AI solutions, software transparency as a quality requirement focuses on information transparency (show what data is used) and process transparency (explain how it works and how results are generated) (Zinovatna and Cysneiros, 2017). This correlates with the companies' ethical guidelines on the transparency of being open and explainable.

Achieving ***trust*** in AI solutions is the goal for the companies as it is directly related to their company value. From the perspective of the interviewees', it is predicted that following the ethical guidelines consistently helps in gaining users' trust. That is the reason why trust is associated with guidelines relating to transparency, openness, privacy, and fairness. In the literature study on ethical issues of AI, trust emerges as a core concept which had its association with ethical issues directly or indirectly (Jones et al., 2015; Pieters, 2011; Elia, 2009). When the ethical issues are regulated it leads to the increased reliability of AI solutions. So, trust plays a vital role when designing AI solutions. Transparency and trust are recommended as quality requirements which support one another (Cysneiros et al., 2018).

So, the proposal made here is, ethical guidelines of AI solutions can be used to define the quality requirements of AI solutions. The ethical guidelines accountability, privacy, security, safety, sustainability, transparency and trust are relating to the quality requirements found in the existing

literature (Section 4.3). Privacy, security and safety are already well-established quality requirements. Sustainability is presented as one of the QRs in Mairiza et al.'s (2010) study. But relating the sustainability ethical guideline with quality requirement needs further detailed investigation. Accountability, transparency and trust are suggested as quality requirements of AI solutions (Cysneiros et al., 2018). The ethical guideline relating to trust does not relate to the characteristics of the system.

The new candidates explainability and fairness are popular ethical guidelines in the case companies. Explainability and fairness exhibits potential to be quality requirements when designing AI solutions. Therefore, accountability, explainability, fairness, and transparency are quality requirements that are specific to AI solutions. The ethical guidelines in the case companies can be used in prioritising the quality requirements depending on AI solutions being developed.

Ethical practices and RE practices

Apart from the ethical guidelines that relate to ethical issues, the interviewees highlighted the usage of multi-disciplinary development teams when designing and developing AI solutions. It was mentioned that, when including people of different skillsets and background, ethical concerns are identified and handled from various perspectives. The results of this case study also highlight an ethical guideline to define the purpose and impact of the AI solutions. It exhibits the relation to one of the core tasks of defining the purpose of the system during RE activities.

5.4 Limitations of the study

The multiple-case study approach has its constraints and Yin's (2013) tests are used to validate the quality of the study. The concerns which are important in our study are reliability, construct validity and external validity.

Reliability

As one of the researchers does not speak Finnish, there was a chance that the interviewee might have missed telling information because of the language barrier (Yin, 2013). This threat was countered by taking the following actions: 1) By making sure that one of the interviewees is a native Finn; 2) the interviewees are fluent in English; and 3) encouraging interviewees to switch language whenever during the interview to avoid loss of information which might lead to a lack of reliability in the results. To avoid the bias that researchers could pass to the interview study because of their work experience and profuse interest, the following three guidelines were followed (Alsaqaf et al., 2018): 1) the interview questions were reviewed by senior researchers and by an experienced researcher who was not involved in this study; 2) conducting a pilot study to measure the feasibility of interview questions, and 3) by audio recording and transcribing the interviews.

Construct validity

To alleviate the threats of construct validity, we employed the following action in the case studies (Yin, 2013; Alsaqaf et al., 2018): 1) we ensured to utilize 'multiple sources of evidence'; 2) member-checking (Yin, 2013; Alsaqaf et al., 2018) during the interview, the interviewing researcher summarized the interviewee's answer and asked follow-up questions to confirm the level of accuracy; 3) Documentation of interview notes to recognize the researchers' initial inference from the interview; 4) by following Yin's tactics to increase construct validity (Yin, 2013) a draft report was created of the case study which was reviewed to identify any interviewing researchers' bias on data collected.

External validity

It is important to know whether the results and findings from the study are generalizable (Yin, 2013). Would the ethical guidelines used in the case companies be adopted by companies in other countries? Could the ethical guidelines from case companies be applied to a company of a similar domain? We can only claim, that these ethical guidelines are proposed based on different ethical issues which might be arising elsewhere too i.e. in any companies designing AI solutions. But there is a risk on implementation in any companies because there could be dissimilarities on the company's contexts. Moreover, the ethical guidelines for AI are generic and high-level, there can be modification when they are applied in real-time projects. Therefore, the list of ethical guidelines for the ethical issues are based on the current results that may be subjected to change based on future case studies. As the study on this area is in its early stage, the results of this thesis are important. To apply the guidelines in real-time projects, the traditional software development process might need a few additional practices which undoubtedly requires more research.

6 Conclusions

The research problem that this thesis attempted to answer was stated as follows: *How can companies solve potential ethical issues of AI solutions during RE activities?* It sets out with the analysis of different ethical issues of AI solutions, and the quality requirements for building AI solutions that were suggested in the literature. The empirical study attempted to discover the ethical guidelines for AI solutions that the case companies have devised to solve ethical issues. Then the relation between ethical guidelines and quality requirements of AI solutions are elaborated in the empirical study summary. The results indicate the role of ethical guidelines when defining the quality requirements of AI solutions during RE activities. The following three statements are the prime conclusions of this thesis that provide answer for the research problem.

The results of this thesis suggest that companies can use ethical guidelines of AI solutions for identifying and prioritising the essential quality requirements of AI solutions.

The case companies of this thesis have defined ethical guidelines to handle ethical issues when building AI solutions. The ethical guidelines developed in the companies relates to accountability, correctness, explainability, fairness, openness, privacy, safety, security, responsibility, sustainability, transparency and trust. The ethical guideline related to accountability focusses on mitigating risks by having accountable person(s) to handle and resolve ethical concerns when developing AI solutions. Likewise, ethical guidelines relating to transparency are about operating openly and making understandable decisions from users' perspective. Making no compromises when performing analytics with user data and operating openly and transparently are emphasized in ethical guidelines relating to trust.

The literature study of this thesis also points out that accountability, privacy, safety, security, transparency and trust as ethical issues of AI solutions. In order to address these ethical issues when building autonomous systems, RE researchers have suggested accountability, privacy, security, safety, transparency and trust as quality requirements. This signifies the potential of utilising the ethical guidelines developed in the companies to identify and prioritise the important quality requirements of AI solutions.

The results of the thesis indicate that accountability, explainability, fairness and transparency are specific quality requirements of AI solutions.

Explainability is closely related with transparency and trust. The issue with AI solutions is that it is hard even for the people who developed the solution to explain how and why a result or recommendation is made. The ethical guideline on fairness focuses on avoiding discrimination and bias when forming results with the data. All the three case companies have ethical guideline relating to fairness. The purpose of this guideline is to avoid to avoid discrimination when utilising data. This signifies the importance and relevance of fairness when developing AI solutions. Therefore,

explainability and fairness highlighted by the case companies can be the specific quality requirements of AI solutions.

Accountability and transparency are featured as quality requirements of AI solutions in existing RE literature. The case companies have also devised ethical guidelines relating to accountability and transparency for developing AI solutions. So, existing literature and the case companies recognized the importance of accountability and transparency when developing AI solutions. As emphasized by the case companies the ethical guideline relating to explainability and fairness are essential for AI solutions. Thus, accountability, explainability, fairness and transparency are recognised as the quality requirements which are specific to AI solutions.

Companies can use multi-disciplinary development teams to define critical quality requirements of AI solutions during RE activities.

It is important to utilise multi-disciplinary teams during the AI solutions development. This ensures in defining the critical quality requirements from diversified perspectives. The case companies emphasize to have teams with different skillset, capabilities and background. The multi-disciplinary teams include experts like data scientists, designers, and sociologists. The quality requirements of AI solutions are defined to identify and eliminate ethical issues and ensure quality of AI solutions. Identifying and defining critical quality requirements of AI solutions assist in achieving quality AI solutions.

The results also indicate that multi-disciplinary development teams approach help in assessing the impact of AI solutions. This practice of utilising multi-disciplinary teams confirms quality of the developed AI solutions from diversified viewpoint throughout the AI solution development cycle.

Scope for future research

The work of Aydemir and Dalpiaz (2018) on ethics-aware software engineering foregrounded the use of ethics requirements in RE process. The ethics requirements are portrayed as ethical values that companies strive for. The ethical issues and misuse cases with the software can be used to identify ethics requirements (Aydemir and Dalpiaz, 2018). Likewise, we see the potential of using ethical guidelines when defining ethics requirements. In the future, we will analyse and clarify the relationship between ethics requirements and quality requirements for ensuring the quality of the AI solutions.

Furthermore, the European Commission has published Ethics Guidelines for Trustworthy AI in April 2019 (European commission, 2019). The guidelines include four ethical principles which are: Respect for human autonomy, Prevention of harm, Fairness, and Explicability (European commission, 2019). In addition, the guidelines present seven key requirements that have to be accounted throughout the lifecycle of AI systems. The key requirements are (European commission, 2019):

1. Human agency and oversight
2. Technical robustness and *safety*
3. *Privacy* and data governance
4. *Transparency*
5. Diversity, non-discrimination and *fairness*
6. Societal and environmental wellbeing
7. *Accountability*

These seven key requirements are in line with the ethical issue identified in this study and support the findings. The next step is to create best practices for implementing these guidelines and requirements in practice.

In Finland, the importance of ethics of AI solutions has recently gained a lot of interests among practitioners and researchers. The Ministry of Economic Affairs of Finland organized an 'Ethics Challenge' in 2018 which challenged companies to develop AI ethical guidelines. Around 70 companies in Finland participated in this challenge. An important direction of our future research is to broaden this case study by exploring the kind of ethical guidelines for AI solutions defined by other companies. We have also planned to do the case studies to investigate how these ethical guidelines are used in AI projects. In addition, we need to analyse the Ethics Guidelines for Trustworthy AI proposed by European Commission and compare it with the case companies' guidelines.

References

Allen, P.J. and Roberts, L.D. (2010), "The Ethics of Outsourcing Online Survey Research," *International Journal of Technoethics,* Vol. 1, no.3, pp. 35-48.

Almquist, E., Senior, J., and Bloch, N. (2016) "The Elements of Value: Measuring and delivering what consumers really want", *Harvard Business Review*, Vol.94, no.9, pp.47-53

Alsaqaf, W., Daneva, M. and Wieringa, R. (2018) "Understanding challenging situations in agile quality requirements engineering and their solution strategies: insights from a case study," *RE2018*, pp. 274-285.

Ameller, D., Galster, M., Avgeriou, P., and Franch, X. (2016), "A survey on quality attributes in service-based systems", *Software Quality journal,* Vol. 24, pp.271-299.

Arnold, T. and Scheutz, M. (2018), "The "big red-button" is too late: an alternative model for ethical evaluation of AI systems," *Ethics and Information Technology,* Vol.20, no.1, pp.59-69.

Aurum, A. and Wohlin, C. (2003), "The fundamental nature of requirements engineering activities as a decision-making process", *Information and Software Technology*, Vol.45, no.14, pp. 945-954.

Aydemir, F.B. and Dalpiaz, F. (2018), "A Roadmap for Ethics-Aware Software Engineering," *ACM/IEEE International Workshop on Software Fairness (FairWare'18),* pp.15-21.

Bonnemains, V., Claire, S., and Tessier, C. (2018), "Embedded ethics some technical and ethical challenges," *Ethics and Information Technology,* Vol.20, no.1, pp.41-58.

Bostrom, N. and Yudkowsky, E. (2011), *The Ethics of Artificial Intelligence,* Cambridge Handbook of Artificial Intelligence, Cambridge University Press.

Boyce, C. and Neale, P. (2006), "Conducting in-depth interviews: A Guide for designing and conducting in-depth interviews," *Evaluation,* Vol.2, no. May, pp. 1-16.

Cassell, C. and Symon, G. (2012), *Essential Guide to Qualitative Methods in Organizational Research,* SAGE publication.

Charmaz, K. (2006), *Constructing Grounded theory: a practical guide through qualitative analysis,* SAGE publications.

Conklin, J. (2003), "Dialog Mapping: Reflections on an Industrial Strength Case Study," *Vis. Argumentation,* pp.1-15.

Chung, L. and Leite, J.C.S.P. (2009), "On Non-Functional Requirements in Software Engineering," *Springer-verlag Berlin Heidelberg,* pp.363-379.

Cysneiros, L.M. and Leite, J.C.S.P. (2004), "Non-functional requirements: From elicitation to Concept models", *IEEE Transactions on Software Engineering*, vol.30, no.5, pp.328-349.

Cysneiros, L.M., Raffi, M.A., and Leite, J.C.S.P. (2018), "Software Transparency as a Key Requirement for Self-Driving Cars," *RE2018*, pp. 382-387.

Dignum, V. (2018), "Ethics in artificial intelligence: introduction to the special issue," *Ethics and Information Technology,* Vol.20, no.1, pp.1-3.

Doyle, T. and Veranas, J. (2014), "Public anonymity and the connected world," *Ethics and Information Technology,* Vol.16, pp.207-218.

Duboc, L., Betz, S., Penzenstadler, B., Kocak, S.A., Chitchyan, R. Leifler, O., Porras, J. Seyff, N. and Venters, C.C. (2019), "Do we really know what we are building? Raising awareness of potential Sustainability Effects of Software Systems in Requirements Engineering", *IEEE Requirements Engineering Conference'19.*

Ebert, C. (1998), "Putting requirements management into praxis: dealing with the non-functional requirements", *Information and Software Technology,* Vol.40, no.3, pp.175-185.

Eisenhardt, K.M. and Graebner, M.E. (2007), "Theory building from cases: Opportunities and Challenges," *Academy of management journal,* Vol. 50, no.1, pp.25-32.

Eisenhardt, K.M. (1989), "Building Theories from Case Study Research", *Academy of management review,* Vol. 14, no.4, pp. 532-550.

Elia, J. (2009), "Transparency rights, technology, and trust," *Ethics and Information Technology,* Vol.11, pp.145-153.

Etzioni, A. and Etzioni, O. (2016), "AI assisted ethics", *Ethics and Information Technology,* Vol.18, no.2, pp.149-156.

European Commission, *Ethics Guidelines for Trustworthy AI,* https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines, retrieved April 9, 2019.

Floridi, L. (2001), "Information Ethics: An Environmental Approach for Digital Divide," *Philosophy in the contemporary world,* Vol 9, no.1, pp.1-7.

Glinz, M. (2007) "On Non-Functional Requirements", *IEEE International Requirements Engineering Conference*, pp. 21-26.

Hosseini, M., Shahri, A., Phalp, K., and Ali, R. (2018), "Four reference models for transparency requirements in information system," *Requirements Engineering,* Vol.23, no.2, pp. 251-275.

Huang, J., Settimi, R., Zou, X. and Solc, P. (2006), "The Detection and Classification of Non-Functional Requirements with Application to Early aspects", *IEEE,* pp. 39-48.

Iqbal, J. and Beigh, B.M. (2017), "Computer Ethics from Obscure to Ubiquitous," *International Journal of Advance Research in Computer Science,* Vol.8, no.3, pp.983-990.

Jones, S., Hara, S., and Augusto, J.C. (2015), "eFRIEND: an ethical framework for intelligent environment development," *Ethics and Information Technology,* Vol.17, pp.11-25.

Karlsson, L., Dahlstedt, A.G., Regnell, B., Natt och Dag, J., and Persson, A. (2007), "Requirements engineering challenges in market-driven software development – an interview study with practitioners", *Information and Software Technology,* Vol.49, no.6, pp.588-604.

Kotonya, G. and Sommerville, I. (1998), *Requirements Engineering – Processes and Techniques*. John Wiley & Sons.

Lauesen, S. (2002), *Software Requirements -Styles and Techniques,* Addison-Wesley.

Loucopulos, P. and Karakostas, V. (1995), *System Requirements Engineering,* McGraw Hill Book Company Europe.

Leite, J.C.S.P. and Cappelli, C. (2010), "Software Transparency," *Business &Information System Engineering,* Vol.2, no.3, pp.127-139.

Mahmoud, A. and Williams, G. (2016), "Detecting, classifying, and tracing non-functional software requirements", *Requirements Engineering*, Vol.21, pp. 357-381.

Mairiza, D., Zowghi, D., and Nurmuliani, N. (2010), "An Investigation into the Notion of Non-Functional Requirements", *ACM Symposium on Applied Computing (SAC),* pp.311-317.

Mylopoulos, J., Chung, L. and Nixon, B. (1992), "Representing and Using Non-Functional Requirements: A Process Oriented Approach", *IEEE Transactions on Software Engineering,* Vol. 18, no.6, pp. 483- 497.

Nuseibeh, B. and Easterbrook, S. (2000), "Requirements Engineering: A Roadmap," *Proceedings of the Conference on The Future of Software Engineering*, pp. 35–46.

Oxford Internet Institute, *Digital Ethics Lab,* [Online] Available: https://www.oii.ox.ac.uk/research/digital-ethics-lab/

Paech, B. and Kerkow, D. (2004), "Non-Functional Requirements Engineering - Quality is Essential" *In: 10th Anniversary International Workshop on Requirements Engineering: Foundation for Software Quality*, REFSQ 2004.

Paetsch, F., Eberlein, A. and Maurer, F. (2003), "Requirements Engineering and Agile Software Development", *IEEE computer society,* pp. 1-6.

Palm, E. (2009), "Securing privacy at work: the importance of contextualized consent", *Ethics and Information Technology,* Vol.11, pp.233-241.

Philips, L.B., Aurum, A., and Svensson, B.R. (2012), "Managing Software Quality Requirements", *Software Engineering and Advanced Applications,* pp.349-356.

Pieters, W. (2011), "Explanation and trust: what to tell the user in security and AI?" *Ethics and Information Technology,* Vol.13, pp.53-64.

Poort E.R., Martens, M., Weerd, I.V.D., and Vliet H.V. (2012), "How Architects See Non-Functional Requirements: Beware of Modifiability", *REFSQ 2012,* pp.37-51.

Quinn, M. (2014), *Ethics for the information age,* Pearson Education Inc.,

Rahwan, I. (2018), "Society-in-the-loop: programming the algorithmic social contract," *Ethics and Information Technology,* Vol.20, no.1, pp.5-14.

Ross, D. and Schoman, K. (1977), "Structured Analysis for Requirements Definition", *IEEE Transactions on Software Engineering,* Vol.3, no.1, pp. 6-15.

Royakkers, L., Timmer, J., Kool, L., and Est, R.V. (2018), "Societal and ethical issues of digitization," *Ethics and Information Technology,* Vol.20, pp.127-142.

Shetty, S. (2017), "Getting Digital Ethics Right,". [Online] Available: https://www.gartner.com/smarterwithgartner/getting-digital-ethics-right/

Sommerville, I. (2010), *Software Engineering, 9th edition*, Nine. New York: Addison-Wesley.

Sommerville, I. and Sawyer, P. (2003), *Requirements engineering: a good practice guide*, John Wiley & Sons, Inc.

Spinello, R. (2012) "Information and Computer Ethics: A brief history," *Journal of Information Ethics,* Vol.21**,** no.2, pp. 17-32.

Stanford Encyclopedia of Philosophy (2001), "Computer Ethics: Basic Concepts and Historical Overview,". [Online] Available: https://plato.stanford.edu/archives/win2001/entries/ethics-computer/

Svensson, B.R., Gorschek, T., and Regnell, B. (2009), "Quality Requirements in Practice: An Interview Study in Requirements Engineering for Embedded Systems", *Requirements engineering: Foundation for Software Quality,* pp.218-232.

Svensson, B.R., Höst, M. and Regnell, B. (2010), "Managing quality requirements: A systematic review", In *Proceedings of the 36th EUROMICRO conference on Software Engineering and Advanced Applications,* pp.261-268.

Svensson, B.R. (2011), "Supporting Release Planning of Quality Requirements: The Quality Performance Model", *Doctoral Dissertation,* Lund University, Department of Computer Science, Lund, p.262

Svensson, B.R., Gorschek, T., Regnell, B., Torkar, R., Shahrokni, A., and Feldt, R. (2012), "Quality Requirements in Industrial Practice – An Extended Interview Study at Eleven Companies", *IEEE Transactions on Software Engineering,* Vol.38, no.4, pp.923-935.

Thayer, R. and Dorfman, M. (1990), *Systems and Software Requirements Engineering,* IEEE Computer Society.

Vampley, P., Dazeley, R., Foale, C., Firmin, S., and Mummery, J. (2018), "Human-aligned artificial intelligence in a multi objective problem," *Ethics and Information Technology,* Vol.20, no.1, pp.27-40.

Yin, R.K. (2013), *Case Study Research Design and Methods,* Sage.

Zinovatna, O. and Cysneiros, L.M. (2015), "Reusing Knowledge on Delivering Privacy and Transparency Together," *in Proc. Of 5th Int. Work. On Requirements Patterns (RePa 2015).*

Zave, P. (1997), "Classification of research efforts in requirements engineering", *ACM Computing Surveys (CSUR)*, Vol.29, no.4, pp. 315-321.

Appendix

Appendix 1

Interview Plan

Objective of the interview

Goal

The goal of this interview is to explore, how organizations see ethics and ethical issues when developing services. The impact made by various AI technologies and applications are the motivation to analyse, how organizations view and handle ethical issues. Uncovering this information will let us know the role of ethics in organization and assists in proposing appropriate practices or suggestions during service development. We can also learn whether there are any existing ethical practices in organizations.

Duration of interview 60-90 mins

Background questions

1. What is your role in the organization?
2. What is your background?

Questions related to ethics

1. Do you have any discussion about ethics in your company?
    a. How about related to digital service development?
    b. How about in a project?
    c. Have you heard about ‘digital ethics’ or ‘data ethics’?
2. What is your view on ethical issues related to digitalization? (Is that changing the situation anyway)
3. Have you faced any ethical questions or ethical issues related to new digital services?
    a. Have you had any ethical questions or issues during service development?
        i. Can you give some examples of ethical issues?
    b. How did you handle those ethical questions or ethical issues?
4. Does your company have an ethical approach or guidelines for digital service development?
    a. Can you tell me about the ethical approach/guidelines?

b. Have you utilized the approach or guidelines in service development?
    i. What is your opinion about it?
c. What is the experience of using that ethical approach/guidelines?

5. Is there a need for common ethical guidelines for service development in your company?
    a. In your point of view, how ethics related to digital service development could be improved or developed in your company?

** How do you see the connection/relation between business ethics and digital ethics? **

Appendix 2

Table 14 is the overall summary of ethical guidelines of AI solutions present in the documents of the case companies. The table is formed based on the document analysis.

Table 14: Ethical guidelines for the use of AI in case companies (From documents)

Company A	Company B	Company C
People-oriented approach -Deploy AI responsibly and promote wellbeing of customers. **Openness and transparency** -Operate openly in relation to our customers, partners and stakeholders and ensure transparency of AI application. **Impact assessment** -Examine carefully how the choices affect the customers and their environment. **Ownership** -Assign owners to all the principles and algorithms that we develop. **Privacy protection** -Safeguard the protection of privacy and personal data in accordance with data protection policies.	**Responsibility and Security** -Protect data and inform users transparently on how and where the data have been utilised. -Responsibility and security direct the collection and utilisation of data. **Customers' needs first** -Create solutions that are useful for our solutions. -Customers decide themselves what data they provide for our use. **Human touch and quality** -Humans lead and develop the AI solutions and algorithms, so we are responsible for the decisions they make. **Common good for society** -Use AI solutions to build better society and better world. -Respect human rights and the AI solutions must not lead to discrimination. **Continuous development of principles** -Constantly review principles and ways of operating.	**Purpose and Impact** -Focus on purpose and impact over technology. **Transparency and Trust** -Prioritize transparency in the system that we design and build and strive to increase trust in all of them. **Inclusion and Fairness** -Aim for inclusion by striving to understand who are all the people that will be impacted by the system that we design and build. **Privacy and Safety** -Collect, store and use personal data safely and default to high privacy.